— (12) 001597-4-T

AD A021 497

social science research institute

USC

UNIVERSITY OF SOUTHERN CALIFORNIA

TECHNICAL REPORT

AN OVERVIEW, INTEGRATION, AND EVALUATION OF UTILITY THEORY FOR DECISION ANALYSIS

DETLOF VON WINTERFELDT

SPONSORED BY:
ADVANCED RESEARCH PROJECTS AGENCY
DEPARTMENT OF DEFENSE

MONITORED BY:
ENGINEERING PSYCHOLOGY PROGRAMS
OFFICE OF NAVAL RESEARCH
CONTRACT NO. N00014-75-C-0487, ARPA
ORDER #2105

AUGUST 1975

SSRI RESEARCH REPORT 75-9

Social Science Research Institute
University of Southern California
Los Angeles, California 90007
213-746-6955

The Social Science Research Institute of the University of Southern California was founded on July 1, 1972 to permit USC scientists to bring their scientific and technological skills to bear on social and public policy problems. Its staff members include faculty and graduate students from many of the Departments and Schools of the University.

SSRI's research activities, supported in part from University funds and in part by various sponsors range from extremely basic to relatively applied. Most SSRI projects mix both kinds of goals — that is, they contribute to fundamental knowledge in the field of a social problem, and in doing so, help to cope with that problem. Typically, SSRI programs are interdisciplinary, drawing not only on its own staff but on the talents of others within the USC community. Each continuing program is composed of several projects; these change from time to time depending on staff and sponsor interest.

At present (Spring, 1975), SSRI has four programs:

Criminal justice and juvenile delinquency. Typical projects include studies of the effect of diversion on recidivism among Los Angeles area juvenile delinquents, and evaluation of the effects of decriminalization of status offenders.

Decision analysis and social program evaluation. Typical projects include study of elicitation methods for continuous probability distributions and development of an evaluation technology for California Coastal Commission decision-making.

Program for data research. A typical project is examination of small-area crime statistics for planning and evaluation of innovations in California crime prevention programs.

Models for social phenomena. Typical projects include differential-equation models of international relations transactions and models of population flows.

SSRI anticipates continuing these four programs and adding new staff and new programs from time to time. For further information, publications, etc., write or phone the Director, Professor Ward Edwards, at the address given above.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (14) USC-001597-4-T, SSRI-RR-75-9	2. GOVT ACCESSION NO. (9)
3. RECIPIENT'S CATALOG NUMBER: Technical research rept.	
4. TITLE (and Subtitle): (6) An Overview, Integration, and Evaluation of Utility Theory for Decision Analysis	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: none
7. AUTHOR(s): (10) Detlof von Winterfeldt	8. CONTRACT OR GRANT NUMBER(s): (15) N00014-75-C-0487, ARPA Order-2105
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Social Science Research Institute, University of Southern California, Los Angeles, California 90007	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (16) NR-197-021, ARPA Order No. 2105
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209	12. REPORT DATE: (11) 1 Aug[redacted]975
	13. NUMBER OF PAGES: 80
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Engineering Psychology Programs, Office of Naval Research, Arlington, Virginia 22217 (12) 85p.	15. SECURITY CLASS. (of this report): unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

utility models
measurement theory
risk
decision analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report is a survey of the measurement theoretic literature on utility models and assessment. Its purpose is to communicate the concepts of measurement theory to decision analysts who may benefit from the application of measurement theoretic models and methods when solving real world evaluation problems. The report is, first, an inventory and dictionary that classifies, translates, and integrates existing measurement theories; and second, an evaluation of the usefulness of measurement theory as a tool for solving complex decision problems. (over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF 014-6601

392 066

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

mt

(Cont.)

The first part of the report explains the role of utility theory as a part of the general theory of measurement, and it develops a classification scheme for utility models. Models are classified according to some of their formal properties and according to the type of decision situations to which they apply. The second part of the report describes the main utility models -- weak order measurement, difference measurement, bisymmetric measurement, conjoint measurement, and expected utility measurement -- through their assumptions, model forms, formally justified assessment procedures, and common approximation methods. In the third part some similarities and differences among models and assessment procedures are discussed. Topics include logical relationships between moels, similarities in the cognitive processes involved in different assessment procedures, and model convergence by insensitivity. The fourth and final part of the report evaluates the use of utility theory for decision analysis, as a tool in formal treatments of decision problems. This analysis concludes that utility theory can be quite useful in structuring evaluation problems, and in eliciting appropriate model forms, but the theoretically feasible assessment procedures are often too clumsy and complicated to be applicable in real world preference assessment. A general critique of current trends to mathematize utility theory concludes the report.

An Overview, Integration, and Evaluation
of Utility Theory for Decision Analysis

Technical Report
1 August 1975

Detlof von Winterfeldt

Social Science Research Institute
University of Southern California

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Engineering Psychology Programs, Office of Naval Research under Contract No. N00014-75-C-0487, ARPA, Order #2105.

Approved for Public Release;
Distribution Unlimited

SSRI Research Report 75-9

Table of Contents

Introduction

This report is a survey of the already voluminous and fast-growing measurement theoretic literature on utility modeling and assessment. It is written specifically for decision analysts who are interested in the use of these abstract measurement theories for solving complex real world decision problems. The main purpose of the report is to connect current theory of utility measurement with decision analytic practice.

Presently, a gap exists between theory and practice, partly because utility theories are formulated in a highly mathematical language that is difficult to relate to real decision problems and real preferences. Many theoreticians overemphasize the mathematical elegance of utility modeling and assessment and show little concern about model applicability. Easy translations and tutorials exist only for a few classes of utility models; the bulk of measurement theories, on the other hand, is hidden in mathematical journals and books. Consequently, many decision analysts who could apply utility theory as a tool for solving complex decision problems find the utility theory literature inaccessible and little use is made of the wealth of models and assessment procedures that utility theory offers.

This report tries to bridge the gap between the theory and practice of utility measurement by:

1. Providing a classification, translation, and integration of utility theories that should make them accessible to the less mathematically sophisticated decision analytic practitioner; and

2. Evaluating the usefulness of utility theory for decision analytic modeling and assessment in order to articulate the needs and considerations of the practitioner for the theoretician.

With these two tasks this review assumes a rather peculiar position among the approximately 20 review articles and books on utility theory that have appeared since the late 60's. It clearly is not a mathematical review as, for example, the books and articles by Luce and Suppes (1967), Fishburn (1970) and Krantz, Luce, Suppes, and Tversky (1971). Neither is it meant to be a tutorial in the application of utility theory such as the books by

Raiffa (1968), Schlaifer (1969), Brown, Peterson, and Kahr (1974), and Keeney and Raiffa (1975). And it does not simply seek to describe current models and assessment procedures for decision analysis as, for example, the reviews by Fishburn (1967), Huber (1974), MacCrimmon (1973), and Kneppreth, Gustafson, Leifer, and Johnson (1974).

Instead, the report hopes to provide the decision analytic practitioner an intelligible and yet comprehensive perspective of utility theory and an overview of the state of the art. It tries to answer questions like these: What utility models are presently available? Where can one read in detail about them? What are the basic characteristics of the models and the assessment procedures? What are the integrating factors? And finally, the report addresses the all important question: How relevant is all this theorizing to the practitioner?

To answer these questions, the report is organized as follows. The first part discusses some general aspects of utility theory as part of measurement theory and it develops a classification scheme for utility models. In the second part, the main model classes (weak order measurement, difference measurement, bisymmetric measurement, conjoint measurement, and expected utility measurement) are described through their assumptions, model forms, formally justified assessment procedures, and approximation methods. The third section of the report looks at some similarities and differences between models and assessment procedures. Topics are the logical relationships between models, similarities and differences in the cognitive processes involved in different assessment procedures, and model convergence by insensitivity. The fourth and final part of the report evaluates the use of utility theory as a practical tool in formal treatments of decision problems. The use of utility theory in structuring evaluation problems and in eliciting appropriate model forms is considered as well as the use of utility theory in scaling and assessment. The report concludes with some general remarks about current trends in utility theory and their implications for the use of utility theory.

Measurement theory and utility models

<u>What is utility theory</u>? -- Utility theory is a part of measurement theory that deals with evaluating (indexing) valuable objects by numbers that are consistent with the decision maker's (group's, organization's) preferences, tastes and values. Utility theory is a collection of models and evaluation procedures that differ in <u>what</u> they measure (e.g., gambles, investment plans, cars), <u>how</u> they measure it (e.g., by adding, by taking expectation, etc.), <u>for whom</u> the measurement is performed (e.g., for an individual, a group, or an organization), and <u>for what purpose</u> the objects are to be measured (e.g., to describe an individual's evaluations, to prescribe his decisions, etc.)

Before going into a more detailed discussion of utility theory, it is useful to back up a little and look at the measurement theoretic framework of which utility theory is a part. In measurement theory, subsystems of the number system with their numerical relations and operations are models for real world objects, their relations, and operations. Measurement theoretic models formulate the principles that <u>justify</u> numerical measurement of these objects, and they provide procedures to <u>construct</u> actual scales.

H.v. Helmholtz (1887) was one of the first measurement theorists who considered the problem of measurement as a problem of modeling empirical systems with systems of numbers. His rudimentary measurement postulates were straight generalizations from the axioms of algebra. In a sense, v. Helmholtz required objects to behave like numbers -- otherwise, he would not consider them measurable. But if they behaved like numbers, one could count, add, and subtract them like numbers, as well as comparing their size. Thus one could construct a scale, and the numbers assigned to the objects would behave just like the objects themselves. Unfortunately, the domain of objects that has the properties required by v. Helmholtz's postulates is very small. Measurement theory would not have reached into areas like color measurement, measurement of probability and utility, or even measurement of temperature, if it had been restricted to empirical systems that obeyed v. Helmholtz's postulates.

But there are two ways to broaden measurement theory. One is to look at other subsystems of numbers as measurement models, possibly without operations such as addition and subtraction. Another one is to relax or reformu-

late v. Helmholtz's postulates into empirical axioms that fit the empirical system better. Hölder's theory of length measurement (1901) was an important step in the latter direction. Hölder formulated conditions on the relations and operations of rods that would allow their numerical measurement. His theory also provided the procedure by which length could be measured, namely by laying off a sequence of rods of equal length against rods of unknown length. Of course, this is exactly the procedure that had been used for hundreds of years.

The other approach to broaden measurement theory by identifying different subsystems of numbers has a relatively recent history. Modern measurement theory (see Suppes and Zinnes, 1963; Krantz et al., 1971) uses the mathematical theory of ordered algebraic structures such as ordered semi-groups, ordered groups, field, rings, etc. (see, for example, Fuchs, 1963; Vinogradov, 1969) to prove the feasibility of measurement and to construct scales. An empirical structure of objects to be measured (e.g., stones), their relations (e.g., stone a "displaces more water" than stone b), and their operations (e.g., stones a and b "displace together as much water" as stone c) is analyzed and assumptions (axioms) are stated that characterize this empirical structure as an algebraic structure with certain nice mathematical properties (e.g., transitivity of the relation "displaces more water", or commutativity of an operation "displace together"). Then a numerical structure is identified, containing a subset of the real numbers, with its usual relations (=, >) and operations (:,+,-,.), that has the same algebraic structure. Finally, a function is constructed that assigns to each element (e.g., a stone) in the empirical structure a number (e.g., volume) such that the relations and operations in both structures coincide. This function is called a homomorphism. <u>Measurement, in short, is the construction of a homomorphism between an empirical and a numerical ordered algebraic structure</u>.

This all sounds rather complicated, but is really based on very simple ideas. Measurement requires the creation of some rule by which numbers are assigned to objects (this actually is Steven's, 1936, somewhat antiquated definition of measurement) and that these numbers behave in accordance with the properties of the objects (their relations and operations). There really

are no limits to this basic idea of measurement. One can invent any funny rule to assign numbers to, say rods, and see whether or not these numbers behave in a way that reflects, say, their length expressed by laying rods off against each other and by connecting them. (Krantz et al., 1971, describe some such "funny rules" for length measurement that actually lead to usable scales, although they are quite different from the length scale we normally use).

This is the framework of utility measurement. Utility theory distinguishes itself from general measurement theory in several aspects:

1. The objects to be measured are objects of cost or value (just as stones are objects of extension or of mass). These objects are called decisions, acts, outcomes, etc. In the following, they will be called "choice entities", or just "objects".

2. The relation between these objects is that of preference, expressed by an individual, group, or organization; their surrogates or representatives, etc.

3. The operation on these objects are not directly definable in terms of external manipulations of the objects (like adding two stones in a water-filled container), but either operations are missing altogether or "operation surrogates" are constructed with the help of a human judge.

These last two distinguishing factors introduce a strong subjective element into utility theory. But utility measurement is different from physical measurement (or any other measurement, for that matter) only in the degree of subjectivity, not in absolute standards. Even length measurement requires human judgment somewhere in the process. The real difference (and the challenge to measurement theorists) is the creation and interpretation of operations that are not so obvious and directly observable as they are in other measurement theories. Conjoint measurement theory, one of the most famous psychological measurement theories, was based on exactly such an invention.

The development of a theory to measure preferences, or to assess utilities of valuable objects, begins with identifying the objects that are to be measured. Then the structure of preferences among these objects (as expressed in individual pair comparisons, for example) is characterized in the form of

normative or descriptive* axioms of preferences allow the identification of the preference structure as an algebraic structure. Utility theory then proceeds to prove that given these axioms, numbers can be assigned to the valuable objects by a function (or rule) that preserves preferences (e.g., the object with the higher utility number is also the more preferred) and reflects the properties of the preference structure (e.g., the difference between two utility numbers reflects the relative strength of preference). The course of the proof provides -- often rather well hidden in the mathematics -- the procedure by which these numbers are assigned to the objects.

The assumptions of utility models fall into three categories:

1. Assumptions that the decision maker can exhibit preferences, and that he does so consistently as if he were maximizing something. These assumptions are often summarized as the "weak order" axiom;

2. Independence assumptions that require preferences among choice entiteis to be independent of certain manipulations of these choice entities. These assumptions are called cancellation, monotonicity, preferential independence, utility independence, and the like;

3. "Technical" assumptions that prohibit abnormalities in preferences. One abnormality is that some choice entity is infinitely desirable ("heaven") or infinitely undesirable ("hell"). "Archimedean" axioms prohibit this from occurring. Another abnormality is that certain choice entities cannot be varied finely enough to produce indifferences with some other fixed choice entities. "Solvability" axioms prohibit such gaps in the set of choice entities.

These assumptions formulate utility theory as a specific model of the decision problem and the decision maker's (group's, organization's) preference structure. These models vary in their formal properties -- particularly in the strength of their assumptions -- and in their interpretation within a specific decision problem, i.e., the model content. There are as many ways to measure utility as there are different types of valuable objects, preference properties, decision makers, etc. These differences in formal model pro-

---

*Utility theory itself is silent about the distinction between normative and descriptive assumptions. Whether a particular theory has normative or descriptive status depends on the interpretation of its axioms.

perties and model content are reflected in the over 50 utility models that now exist. The most important of these models will be classified, described, and integrated in the following sections of this report.

Classification of utility models -- The two dimensions of model variability discussed above will now be used to classify measurement theories. First, utility models will be classified according to some of their formal properties. Then a classification scheme for possible decision situations will be presented.

The following formal distinctions between utility models are made:

1. Deterministic vs. probabilistic models;
2. Ordinal vs. interval models.

Probabilistic models express utility and preferences as a result of a random process. Utilities are assessed by determining "probabilities of preferences", presumably by repeated observations of preferences among valuable objects. Predictions of these models state a probability that an object is chosen over another as a function of their numerical utilities. Deterministic models (also called algebraic models) assume no randomness whatsoever in utilities or preferences. Consequently, both their assessment and their predictions are deterministic, based on a unique set of preferences and indifferences, and on unique predictions. Deterministic models are special cases of probabilistic models, in which only probabilities of 1 and 0 are allowed.

The second distinction refers to the scale quality of the utility function that can be assessed within the framework of a particular model. Ordinal models produce utility functions that make statements about the order of preferences only. The specific shape of these utility functions does not contain any information about the preferences, i.e., utility functions are unique up to a monotone transformation only. Interval models produce utility functions that also make statements about the relative strength of preferences. The shape of these utility functions contains meaningful information about the modeled preferences, but their origin and unit are arbitrary, i.e., they are unique up to a positive linear transformation. Clearly interval models are special cases of ordinal models.

Table 1 presents the main classes of utility models within this simple formal classification scheme. All of these model classes will be dealt with

in more detail later. Also, in later sections of this report, some more formal relationships amont the models in the boxes of Table 1 will be worked out.

Insert Table 1 About here

The model classes in Table 1 can be applied to quite different decision situations which give them their specific interpretation as utility models. The distinguishing characteristic of decision situations is complexity. Decision situations can be classified according to the presence or absence of complexity in the following aspects:

1. static vs. dynamic decision environment;
2. single decision maker vs. multiple decision makers;
3. single aspect choice entity vs. multiple aspect choice entity;
   a. single attributed vs. multi-attributed choice entity;
   b. riskless vs. risky choice entity;
   c. time invariant vs. time variable choice entity;
   d. choice entity that affects only one individual vs. choice entity that affects many.

In static decision situations, decision makers make one decision at a specific time in an unchanging environment; the decision's consequences may reach into the future, however. Dynamic decision situations are characterized by sequential decision making under changing circumstances, changing values, and changing information (see Rapoport, 1975). Decisions in operational management are often highly dynamic, as, for example, dispatching decisions, or inventory control decisions. Strategic decisions, although they usually have long term effects, can often be interpreted as static decisions.

The next important distinction between decision situations addresses the question: utility for whom? A distinction can be made between cases in which a single decision maker evaluates or decides, vs. cases in which a group or a committee has that task. When you evaluate cars for possible purchase, and you finally decide which car to buy, you are the single decision maker, even if you consider the opinions of others and the effects your deci-

Table 1: Main classes of utility models

	deterministic models	probabilistic models
ordinal models	Semiorder measurement (Luce, 1956) Interval order measurement (Fishburn, 1970, 1973a) Lexicographic measurement (Fishburn, 1970, 1974b) Weak order measurement (Krantz et al., 1971)	Weak constant utility model (Luce and Suppes, 1967) Random utility model (Luce and Suppes, 1967)
inteval models	Difference measurement (Krantz et al., 1971) Bisymmetric measurement (Pfanzagl, 1968) Conjoint measurement (Luce and Tukey, 1964; Krantz, 1964) Expected utility measurement (v. Neumann and Morgenstern, 1947; Savage, 1954; Luce and Krantz, 1971)	Strong constant utility model (Luce and Suppes, 1967; Becker et al., 1967) Strict constant utility model (Luce and Suppes, 1967; Becker et al., 1967) Elimination by aspects model (Tversky, 1972a and b)

sion may have on others. Multiple decision makers are involved when a city council evaluates alternative taxation plans, or when a committee adopts a resolution.

The classification aspects 3a-d refer to the question: utility for what entity? The complexity of choice entities can increase in at least four different aspects. (3a) A choice entity is called single attributed if it varies on a single, well defined dimension or attribute. Money and profit rates are single attributed; so are commodities like gasoline, butter, etc. Commodity bundles, cars, social programs, development plans are multiattributed, that is, they vary on several, and often conflicting dimensions of value. Cars, for example, vary on attributes such as cost, comfort, horsepower, cornering ability, etc. In this report, a multiattributed object will often be described as an n-tuple of single attribute values a_i, where $(a_1, a_2, \ldots, a_i, \ldots, a_n)$ denotes a multiattributed object a that has value a_i in the i-th attribute.

(3b) A choice entity is called riskless, if all of its outcomes are determined with certainty. An unconditional monetary gift is riskless. A choice entity is called risky, if some or all of its outcomes are uncertain. Gambles, investment plans, and stocks are risky choice entities. Similar to the n-tuple description of multiattributed choice entities, risky choice entities will often be described as m-tuples of outcomes, $(a_1, a_2, \ldots, a_j, \ldots, a_m)$, where a_j is the outcome to be received if an uncertain event E_j occurs.

(3c) Choice entities are called time invariant, if their consequences are received at a unique time now or in the future. A meal, a car, a site for a plant are time invariant. Choice entities are called time variable if parts or all of their outcomes are distributed over time. Returns from investments are distributed over time; jobs may vary in the prospects for future salary increases, etc. As before, time variable choice entities can be characterized by an N-tuple of outcomes to be received or consumed at different times. $(a_1, a_2, \ldots, a_k, \ldots, a_N)$ would denote a time variable choice entity in which outcome a_k will be received at time t_k.

(3d) Choice entities whose consequences affect a single individual can be distinguished from those that affect many. Individual consumption affects

only the consumer. Public policy decisions affect many. Choice entities that affect many can again be symbolized by an M-tuple $(a_1, a_2, \ldots, a_i, \ldots, a_M)$ where a_i is the consequence for individual I_i. At first glance, this distinction seems to have a lot in common with (2)(individual decision maker vs. group decision maker), and, in fact, they have often been lumped together as a distinction between individual vs. social decision making. However, it is logically possible, and useful for modeling purposes to keep these two distinctions separate. A single decision maker can decide about alternatives that affect many, and a group of decision makers can make choices that affect only one person. Dictators and court juries come to mind as examples.

Whenever choice entities vary on more than one of these complicating aspects, double or triple subscripts will be used to describe their elements symbolically. For example, a choice entity $\underline{a}$ may be uncertain and multiattributed. In this case it would be described by its elements a_{ij} which are single attribute outcomes to be received if the uncertain event E_j occurs.

The above classification scheme leads to 128 distinct decision situations. The 13 models of Table 1 (together with certain model combinations) would generate a huge number of utility models when applied to these 128 decision situations. Naturally, not all models have been applied to all cases. And this report will make some further restrictions on the models and cases that will actually be discussed in detail. These restrictions are discussed below.

Some omissions -- Any honest utility modeling attempt will have to acknowledge the enormous complexity of the decision situation and the inadequacy of strong and simple models. Ideally, one would like to model preferences in a dynamic decision situation in which a group of decision makers makes decisions about choice entities that are multiattributed, risky, time variable, and affect many; and ideally such a model should make the weakest assumptions possible. Practically, one will have to be much more modest, first because models for the most complex decision situations do not exist, and second because weak models usually require an inordinate amount of very complex assessment.

Realizing that the modeling must be simplified one can either simplify the decision situation, or strengthen the model. This report will treat in

detail only utility models that make both simplifications. Eliminated from the discussion are the weakest model categories, (probabilistic models, semi-orders, interval orders, and lexicographic orders) because their practical applicability in real world decision problems is questionable. Furthermore, dynamic decision situations, and group decision makers will be omitted, because appropriate utility models are missing for these cases.

What follows is a brief discussion of the model categories and the decision situations the report leaves out of a detailed analysis. After this discussion, the report will describe the five remaining model categories (weak orders, difference measurement, bisymmetric measurement, conjoint measurement, and expected utility measurement) as they are applied to the remaining decision situations.

Probabilistic models are one of the many possibilities in utility theory to cope with the problem of error and the equivalent problem of fluctuating or changing preferences and/or responses. Probabilistic utility models have built into themselves a theory of random preferences that can account for substantial errors or fluctuations. As a measurement theory these models differ somewhat from the concepts described below. They assume that a numerical assessment of the strength of preference is given by a probability of preference in pairwise comparisons, rankings, or choices. This probability is assumed to be measured through repeated observations of the same comparison or choice, so that some relatively high level of numerical measurement is the base on which these models build. On this base, they formulate conditions of probabilistic preferences -- beyond those of simple probability theory -- that allow the expression of these probabilities as functions of hypothetical utility numbers assigned to the valuable objects. No procedure to estimate preference probabilities or to estimate from these preference probabilities the underlying utilities are given, but statistical estimation methods are available to perform some such estimations, once relative frequencies of preferences are given.

Probabilistic models have been developed for risky and riskless single and multiattributed choice alternatives. They are silent about the remaining classification aspects such as group vs. individual decision making, time

variable or time invariant choice entities, etc. The following probabilistic models have been developed:

1. Constant utility models (see Luce, 1959; Luce and Suppes, 1967);
2. Random utility models (Becker, DeGroot, and Marschak, 1963);
3. Elimination by aspects models (Tversky, 1972a and b).

Constant utility models assume that the randomness of preferences is generated by uncertain responses or an uncertain decision making mechanism. Random utility models hypothesize that this randomness lies in the underlying utilities of the valuable objects themselves. EBA-models assume a sequential probabilistic elimination process, in which, at each stage, one attribute of the choice entities is probabilistically sampled, and all alternatives are eliminated that do not have the attribute (or that are not satisfactory in that attribute). All models express the probability of choice or the probability of preference as a function of numerical utilities. The main assumption behind these models allowing contruction of such functions are stochastic transitivity, simple scalability, etc., all of which are spelled out in detail in the references cited.

The practical impact of probabilistic models on decision analysis has been very small, and in recent years the theoretical development of probabilistic utility models has come to a virtual standstill. Probably the main reason for the lack of use of probabilistic models in decision analysis is the difficulty of practically assessing utility functions. Decision makers usually do not have the time or the patience to carry out the assessment procedures probabilistic models require, and even if time and complexity were not problems, the assumptions of independent repeated responses in the assessment make the results rather dubious. Further objection stems from the weakness of probabilistic models in guiding decisions. Model predictions or prescriptions in form of probabilities of preferences are too weak to be helpful in solving real decision problems.

Difficulty in assessment and weakness in prediction and prescription are also limitations of semiorders (Luce, 1963) and interval orders (Fishburn. 1970). While probabilistic models try to take intransitivities or changing

preferences into account by building a theory of random preferences, these models formulate a weaker basis of consistency. Rather than requiring, as do all other utility models, the decision maker's preferences and indifferences to be transitive, semiorders and interval orders allow intransitive indifferences and require only strict preferences to be transitive. According to these models preferences exist only if choice entities are sufficiently different in utility. Thus they allow indifferences between choice entities, evey if they actually have different utility numbers.

Although these approaches to utility measurement could possibly be used as a theory of approximate measurement in decision analysis (see v. Winterfeldt, 1975), no use has been made of semiorders and interval orders as of the present, presumably because assessment methods within these models are complicated, and predictions and prescriptions are even weaker than those of plain orderings. Also, interval and semiorder models have not yet been extended to more realistic choice entities than those which are simple single attribute, riskless, and time invariant in nature.

Lexicographic models (see Fishburn, 1970, 1974c) apply to multi-aspect choice entities, and they have usually been interpreted as models for multiattribute choice entities. Unlike most utility models, lexicographic models are non-compensatory, that is they do not trade-off one value aspect against another. Instead, they consider each value aspect individually and preferences are determined solely on the basis of that value aspect. Where no preference can be established, the next important value aspect is considered. Situations in which such a non-compensatory model is an appropriate representation of the decision maker's preferences over multi-aspect choice entities are very rare. Although there is evidence that decision makers sometimes use lexicographic orderings as simplifying strategies to determine their preferences (see Tversky, 1969), these strategies can seldom be justified as rational models for decision making. Consequently, lexicographic models have found no application in decision analysis.

This leaves us with a list of five model classes that will be discussed in more detail in this review: weak orders, difference measurement, bisymmetric measurement, conjoint measurement, and expected utility measurement.

In addition, the report will also omit some decision situations, namely dynamic decision situations, and group decisions.

The reason for leaving out dynamic decision situations is very simple: no measurement model deals specifically with the dynamic nature of the decision environment. There exist dynamic programming models which use static expected utility measurement as inputs into their dynamic calculations (see Rapoport, 1967, 1975), but no attempt has been made in the measurement literature to incorporate the dynamic nature of the decision environment into utility measurement as such.

Group decision models or models for social choice have enjoyed increasing attention by measurement theorists through the last few years. But their practical impact on the measurement of group utility functions in decisions analysis is still negligible. Although models for multiple affected individuals have now reached the stage of application (see Kirkwood, 1972; Keeney and Kirkwood, 1973; Keeney, 1975; Keeney and Raiffa, 1975), measurement models for multiple decision makers are still problematic.

Social utility models deal with the following fundamental question: How can individual preferences (or individual rank orders, or individual utilities) be aggregated to a group utility function? Arrow's (1951) famous paradox claims that under some reasonable conditions, no such aggregation rule exists. Since Arrow, several attempts have been made to cope with this problem. One involves changing some of Arrow's conditions in order to resolve the paradox. This literature is best described in Luce and Raiffa (1957) and, more recently, in Fishburn (1973b, 1974b). Other researchers looked at several "reasonable" aggregation rules (voting paradigms) to see how Arrow's paradox actually affects the outcomes (see Fishburn, 1974d) formulated as "voting paradoxes". Rather than providing -- in the spirit of decision analysis -- formal tools to cope with the problem of integrating individual preferences or utilities, most of this research remains critical, full of paradoxes and criticisms of interpersonal utilities, etc. What is lacking are practical characterizations of group decision problems that say: If conditions a-d are fulfilled, this assessment procedure and that aggregation process is feasible.

Recent research by Keeney and Kirkwood moved in this direction by reinterpreting conjoint measurement theory (Krantz, 1964; Luce and Tukey, 1964;

Krantz et al., 1971) and multiattribute expected utility theory (Kirkwood, 1972; Keeney and Raiffa, 1975) for group decision problems. However, these theories are more appropriate for decision problems in which the decision of an individual affects many than for actual group decision making problems. Often, one can think of a group decision making problem as a problem in which a supra decision maker (a term used by Keeney and Raiffa) is created that represents a number of decision makers. Such a supra decision maker would then treat the problem as an individual decision making problem in which the consequences may affect many. This is obviously the sort of paradigm to which our classification aspect 6 is addressed. These models will therefore be discussed in detail later. A later report in this series will deal exclusively with the genuine problem of multiple decision makers.

So much for the omissions of this report. We can now turn our attention to the remaining five model classes as they apply to the remaining decision situations.

The main representations

The 16 decision situations that remain to be discussed characterize the different types of choice entities that were described earlier:

3a. single attributed vs. multiattributed choice entities;
3b. riskless vs. risky choice entities;
3c. time invariant vs. time variable choice entities;
3d. choice entities that affect one individual vs. choice entities that affect many.

These 16 cases vary from choice entities with no complicating aspects (single attributed, riskless, time invariant, choice entities that affect only one person) over choice entities with one complicating aspect (multiattributed <u>or</u> risky <u>or</u> time variable <u>or</u> many individuals affected) to the most complex choice entities that are multiattributed, risky, time variable, and that affect many.

Utility models exist for choice entities with no complicating aspect (Table 2), with one complicating aspect (Table 3), and with two complicating aspects (Table 4). In the case of two complicating aspects, utility models have been developed only for the combination of risky choice entities with

some other complicating aspect.

Insert Tables 2-4 about here

One could, of course, apply models designed for the less complex cases to more complex ones (e.g., weak order model to multiattributed, risky, and time variable choice entities) by ignoring the additional complicating aspect. However, such an approach would lead to models that require extremely complex assessment tasks. Alternatively, one could combine models for each single complicating aspect to an overall model (e.g., a combination of a conjoint measurement model to deal with the multiattribute aspect, an expected utility model to deal with riskiness, and a weak order model to deal with time variability), but such an overall model may ignore interactions between complicating aspects (see v. Winterfeldt and Fischer, 1975). Therefore, the further discussion will be restricted to the five main model classes.

- weak order measurement,
- difference measurement,
- bisymmetric measurement,
- conjoint measurement,
- expected utility measurement

as they apply to the 8 types of choice entities in Tables 2-4 . Since choice entities II-V and VI-VIII have similar structural properties (as suggested by their n-tuple representation on pp. 10-11), the discussion will usually concentrate on the multiattribute case II and the multiattribute risky case VI. Analogous model applications to the time variable cases (III and VII) and to the case in which multiple individuals are affected (V and VIII) will only be sketched.

The five model classes will be discussed under the following topics:

1. Which cases has the model been applied to?
2. What are the main model assumptions?
3. What is the model form?
4. What are the formally justified assessment procedures?
5. What are some reasonable approximation techniques?

Tables 5 and 6 give a preview.

Table 2

Models for choice entities that have no complicating aspect.

(The Roman numbers identify the type of choice entity, the Arabic numbers identify the specific model for that choice entity.)

1. Weak order (Krantz et al., 1971) 2. Difference and ratio measurement (Suppes and Zinnes, 1963; Krantz et al., 1971) 3. Bisymmetric measurement (Pfanzagl, 1968; Krantz et al., 1971)
I

Table 3: Models for choice entities that have one complicating aspect

multiattributed	risky	time variable	multiple individuals
1. Weak order (trade-off) model (Suppes and Zinnes, 1963; Boyd, 1970) 2. Conjoint measurement models (Krantz et al., 1971) a. additive (Luce and Tukey, 1964; Krantz, 1964; Luce, 1966) b. other simple polynomials (Krantz et al., 1971; Tversky, 1967) 3. Bisymmetric measurement models (Fishburn, 1975) a. multilinear 1 b. multilinear 2 c. mutiiplicative d. additive	1. Weak order risk theories (Coombs, 1973; Pollatsek and Tversky, 1970; Huang, 1971) 2. Expected utility (EU) (Fishburn, 1970; Krantz et al., 1971) a. with numerical probability (v. Neumann and Morgenstern, 1947) b. with subjective probability (Savage, 1954) c. for finite utility differences (Davidson, Suppes, and Siegel, 1957) d. conditional expected utility (Luce and Krantz, 1971)	1. Weak order (trade-off) model (Raiffa and Keeney, 1975) 2. Conjoint measurement a. additive, no discounting (Fishburn, 1970) b. additive, variable discounting (Koopmans, 1972) c. additive, constant discounting (Koopmans, 1972)	1. Conjoint measurement a. additive (Keeney and Raiffa, 1975)
II	III	IV	V

Table 4: Models for choice entities that have two complicating aspects.

risky and		
multiattributed	time variable	multiple individuals
1. Weak order-EU-model (Boyd, 1970) 2. Conjoint measurement-EU-models a. additive (Fischer, 1972 Keeney and Raiffa, 1975) b. others (v. Winterfeldt and Fischer, 1975) 3. Bisymmetric-EU-models a. multilinear 1 b. multilinear 2 c. multiplicative d. additive 4. EU-decomposition models a. multilinear 1 (Fishburn, 1973c, 1974a) b. multilinear 2 (Keeney, 1972; Fishburn and Keeney, 1974) c. multiplicative (Keeney, 1973) d. additive (Fishburn, 1965)	1. Weak order-EU-model (Pollard, 1969) 2. Conjoint measurement-EU-models a. additive (Keeney and Raiffa, 1975) b. other (v. Winterfeldt and Fischer, 1975) 3. EU-decomposition models (Keeney and Raiffa, 1975) a. additive, variable discounting (Pollak, 1967) b. multiplicative, constant discounting (Meyer, 1970)	1. EU-decomposition models (Keeney and Raiffa, 1975) a. additive (Kirkwood, 1972; Keeney, 1975) b. multiplicative (Keeney and Kirkwood, 1973)
VI	VII	VIII

Insert Tables 5 and 6 about here

Weak order measurement -- Weak order measurement has been applied to all cases in Tables 2-4 except for the cases of multiple affected individuals. For the risky choice entities, weak orders have been combined with the expected utility assumption to model the non-risk aspect of preferences (i.e., multi-attributed or time variable).

The main model assumption behind weak orders is transitivity of preferences. If the set of choice entities is finite (or even countably infinite), transitivity is necessary and sufficient to prove that a rule (function) can be created that assigns numbers to valuable objects such that the more preferred object has a higher number. In uncountably infinite sets, things become a little more difficult, and some technical assumptions have to be added. The formal weak order representation is:

Weak order representation

$$a \succsim b$$

if and only if

$$u(a) \geq u(b)$$

where a and b are choice entities, "$a \succsim b$" means "b is not preferred to a", u is the rule or function by which numbers are assigned to the choice entities, and u(a) is the utility of a.

Scaling within the weak order model can take 2 forms:

1. Rank ordering;
2. Indifference curve construction.

The first procedure is as simple as measurement can get. In the finite case, the assessor simply rank orders all valuable alternatives, and the rank order number is the utility of a valuable object. Procedures for the infinite case (countable or not) are somewhat more complicated, but they are also based on rankings. The second procedure is applicable in cases where the choice entities have various value aspects (any of the cases II-VII fall under this heading). If the weak order assumption holds, one can construct indifference

Table 5: Model descriptions

The model	Weak order	Difference measurement (algebraic)	Bisymmetric measurement (bisection)
cases the model has been applied to	I, II, III, IV, VI, VIII	I	I, II
main model assumptions	transitivity of preferences	monotonicity, sign reversal	monotonicity, bisymmetry commutativity, idempotency
model form	$a \succsim b$ iff $u(a) \geq u(b)$	$a \succsim b$ iff $u(a) \geq u(b)$ and $ab \succsim cd$ iff $u(a)-u(b) \geq u(c)-u(d)$	$a \succsim b$ iff $u(a) \geq u(b)$ and $u(a \circ b) = \frac{1}{2}u(a)+\frac{1}{2}u(b)$
formally justified assessment	pair comparisons, rankings, indifference curve procedures	standard sequence of equally spaced objects	bisection
approximation methods	-	rating, method of equal appearing intervals, bisection, numerical difference judgments	rating, direct difference judgments, method of equal appearing intervals

Table 6: Model descriptions

The model	Conjoint measurement (additive)	Expected utility measurement
cases the model has been applied to	II, III, IV, V, VI, VIII	III, VI, VII, VIII
main model assumptions	preferential independence	Sure thing, combined archimedean and solvability axiom
model form	$\underline{a} \succsim \underline{b}$ iff $u(\underline{a}) \geq u(\underline{b})$ where $u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i)$	$\underline{a} \succsim \underline{b}$ iff $u(\underline{a}) \geq u(\underline{b})$ where $u(\underline{a}) = \sum_{i=1}^{n} p(E_i)u(a_i)$
formally justified assessment	dual standard sequences	sequence of indifference lotteries
approximation methods	direct assessment of single attribute utilities (rating, bisection, etc.) and weighting	magnitude estimation of probabilities p(E) and indifference lotteries for utilities.

curves (or less graphically, classes of indifferent choice entities), and index such curves by an appropriate numeraire (Raiffa, 1969). These procedures are often quite helpful in decision analysis and some researchers and decision analysts have tried to exploit the weak order assumption alone to construct utility function in complex decision problems (Boyd, 1970; Pollard, 1969).

If one wants to make simplifying assumptions (such as convexity of indifference curves or even linearity of indifference curves) this assessment can be simplified substantially. Sequential application of trade-off procedures can also be used to make the task of constructing indifferences or of comparing choice entities easier (Raiffa, 1969; Keeney and Raiffa, 1975; v. Winterfeldt and Fischer, 1975). Boyd (1970) exploited some of these assumptions to create an interactive technique that finds the best element in a set of choice entities on the basis of local trade-off ratios or substitution rates. MacCrimmon and Toda (1969) and MacCrimmon and Siu (1974) describe interactive techniques to approximate indifference curves.

There is one rather peculiar application of weak order measurement in connection with some much stronger forms of measurement in the risky case III. Several strong theories measure the "riskiness" of uncertain choice entities (see Pollatsek and Tversky, 1970; and Huang, 1971). This measurement of risk in itself, however, does not produce a utility function, but rather a "risk" function that says nothing about preferences. However, a special form of weak order measurement can be applied to measure utility as a function of the riskiness of a gamble and some other aspect of gambles, such as their expected value. In this vein, Coombs (for an excellent summary, see Coombs, 1972) has developed portfolio theory, that can be based on measurement of risk to create a weak order of preferences over gambles varying in riskiness and expected value.

The substantive relation in the risk measurement theories is that an ununcertain choice entity is "perceived to be more risky" than another one. Pollatsek and Tversky (1970) developed a theory of risk measurement that is not unlike Hölder's theory of extensive length measurement. Unlike most utility theories, their theory uses a direct manipulation of gambles, namely that

of convolution (i.e., playing two gambles simultaneously) to define an operation on gambles. This operation is then treated just like the concatenation operation in length measurement that combines two rods. The main assumption in Pollatsek and Tversky's theory is that convoluting two gambles that stand in a certain riskiness relation does not change that relation, if they are both convoluted with the same gamble. For example, if gamble a is more risky than gamble b, and a and b are both played simultaneously with c, then the mixture a <u>and</u> c should still be more risky than the mixture b <u>and</u> c. Together with the usual weak order assumption (this time for the riskiness relation) and appropriate solvability and archimedean axioms the following risk model is implied:

<u>Extensive risk measurement</u>

$$a \overset{\circ}{\succsim} b$$

if and only if

$$R(a) \geq R(b)$$

$$\text{and } R(a \circ b) = R(a) + R(b)$$

where a and b are two risky choice entities "$a \overset{\circ}{\succsim} b$" stands for "b is not perceived to be riskier than a", R is the risk function, and " $\circ$ " stands for the convolution operation.

An alternative to this theory is presented by Huang (1971), who essentially used the v. Neumann and Morgenstern axioms (see p.39) to prove that the expected risk of two gambles preserves the preferences among gambles with risky outcomes. Using the riskiness relation as in Pollatsek and Tversky's theory and the v. Neumann and Morgenstern axioms applied to that relation, the following risk representation can be proven:

<u>Expected risk measurement</u>

$$a \overset{\circ}{\succsim} b$$

if and only if

$$R(a) \geq R(b)$$

$$\text{and } R(apb) = pR(a) + (1-p)R(b)$$

where all symbols have the same meaning as above except that the convolution of gambles is substituted by the symbol apb that denotes a supra gamble which

yields with probability p the gamble a as an outcome, with probability 1-p the gamble b.

To make either of the risk theories a utility theory, one would have to define a function that links perceived risk (numerically measured in R) to preferences (numerically to be measured in utilities). That is, one wants to find a function h such that

$$a \succcurlyeq b$$
$$\text{if and only if}$$
$$u(a) \geqslant u(b)$$
$$\text{where } u(a) = h(R(a))$$

Some restrictions for such a weak order are spelled out in Coombs' portfolio theory (Coombs, 1972).

Construction of the function R depends on the measurement model (extensive or expected risk model). In extensive risk measurement one would use standard sequence procedures, in which a sequence of lotteries is generated by convoluting gambles with identical risks. Arbitrarily assigning a risk of 1 to one gamble and convoluting it with a gamble that has the same riskiness, one would generate a gamble that--by the measurement representation--has a risk of 2. Convoluting this gamble again with a gamble that has equal riskiness as the unit gamble, one would generate a gamble with a risk of 3, etc. In expected risk measurement, risk would be measured by matching the risk of a gamble b that has riskiness between two gambles a and c with a supra-gamble apc by varying the probability p. p then is an index of the riskiness of a. (This "indifference lottery procedure" will later be explained in more detail for preference judgments in expected utility theory.) To construct a utility function over risky choice entities one can then use any of the described weak order procedures to generate a rank order of indifference classes of risky choice entities that are matched in riskiness (have equal R).

The four remaining utility models (difference, bisymmetric, conjoint, and expected utility measurement) are <u>all</u> special cases of the weak order model. Without explicit statement, the weak order model will from now on be assumed to be valid.

<u>Difference measurement</u>* -- Difference measurement is one important way to strengthen utility measurement beyond weak orders. In addition to simple preferences among choice entities difference models also compare the relative difference of the strength of preference between pairs of choice entities. Added to judgments such as "a is preferred to b" are judgments of the form "the difference in strength of preference between a and b is larger than that between c and d". Judgments of this type can be rather difficult, particularly if choice entities are complex. Therefore -- although difference measurement is, in principle, applicable to all cases in Tables 2-4 -- it is reasonable to restrict its discussion to the simplest case I.

Difference measurement is the first modeling approach that uses "operation surrogates". Note that there were no operations whatsoever involved in weak order measurement. In difference measurement one wants to create an operation "addition" of utility differences between choice objects. Somehow, one would like to find two choice entities x and y such that their utility difference equals the "sum" of the utility differences between a and b and c and d. If b=c, then there appears to be an obvious way of defining "addition of judged utility differences"; the sum of utility differences between a and b on one hand and b and c on the other is the judged utility difference between a and c. This idea is really the heart of the "invented" operation. The rest is generalizing this idea to non-adjoining cases.

For example, take the problem of quantifying the degree of displeasure from driving to work as a function of driving time. Obviously, time itself is not a very good measure of that utility cost (or disutility). The extra five minutes added to the one hour ride may create less discomfort than the extra five minutes added to the usual 10 minutes ride. That is, the difference in utility between 65 minutes and 60 minutes is smaller than that between 15 and 10 minutes. Similarly, all differences in time intervals could

---

*Although there are several types of difference measurement models (such as the positive model, the algebraic model, the absolute model, and the conditionally connected models; for details, see Krantz et al., 1971) we will discuss difference measurement here by example of the case that is most typical for utility theory, the algebraic model. This model is also equivalent to a ratio measurement model.

be compared. The operation would then take the following form: the difference in displeasure between driving 5 and 10 minutes, "added to" the difference between driving 10 and 15 minutes, "is equal to" the difference in displeasure between driving 5 and 15 minutes.

The fundamental assumption of difference measurement is that this operation behaves nicely, meaning that adding the same amount of difference to two already established degrees of differences does not alter the relation between the original differences. This is a monotonicity assumption not unlike the usual cancellation property in adding and multiplying numbers. Such independence assumptions are the basis of any higher structured measurement theory. This monotonicity assumption, together with an appropriate sign reversal assumption (if the difference between a and b is greater than that between c and d, then the reverse must be true for the differences between b and a and d and c respectively), and solvability and archimedean axioms produces the following model form:

<u>(Algebraic) difference measurement</u>

$$a \succsim b$$

if and only if

$$u(a) \geq u(b)$$

and

$$ab \;\dot{\succsim}\; cd$$

if and only if

$$u(a)-u(b) \geq u(c)-u(d)$$

where the upper part is the usual weak order representation, and the lower part reads as follows: "$ab \;\dot{\succsim}\; cd$" means "the judged difference between c and d is not greater than the difference between a and b".

The formally justified procedure to assess utility in the framework of difference measurement is to lay out a sequence of choice entities that have equal utility differences and that are connected to one another. This is a type of construction procedure which will come up recurrently in the discussion of utility models and is usually called "standard sequence" because it is a systematic sequence of standard choice entities that are equally spaced

in utility. In the example of driving from and to work, a standard sequence may be constructed by beginning with a small time step from 0 to 5 minutes, and then asking which increase in time from 5 to x would create as much additional discomfort as the increase from 0 to 5 , followed by the same question from x to z, etc. This gives exact utilities for the points which are members of the standard sequence, and approximate utilities for the elements in between. Defining each utility difference to be equal to 1, and the utility of some arbitrary point equal to 0, the utilities of each point in the standard sequence can thus be inferred. The utilities of the intermediate points can be approximated through interpolation, or, alternatively through a finer graded standard sequence (e.g., one that would start with a smaller initial difference).

So much for the formally justified assessment technique. There are numerous scaling procedures which are good approximations of this procedure, not only in the sense that they will yield converging utility functions, but also in the sense that they involve cognitive processes that are similar to those in standard sequences. A method that closely resembles standard sequences is the method of equal appearing intervals (Torgerson, 1958). In this method, two extreme choice entities are given to the assessor (the most and the least preferred one) and he is asked to find a number of intermediate choice entities that subdivide the set into elements of equally appearing utility differences. The method of bisection (Torgerson, 1958; Pfanzagl, 1968) structures this procedure more firmly. In the bisection method, the assessor is asked to determine a choice entity that is equally far in utility from two specified elements. Further subdivision leads to a finely graded scale.

In contrast to these indirect scaling methods, other approximation methods involve direct numerical assessment of choice alternatives. One simple way is to rate utilities directly on a numerical scale (ranging from say 0 to 100). This kind of procedure has been advocated by Edwards (1971) for utility assessment in the multiattribute context. Another procedure requires the decision maker to make direct ratio judgments about the utility (or utility difference) for pairs of choice entities. This procedure has been originally proposed by Stevens (1936) in psychophysical measurement and it was

applied to assess the utility for money by Galanter (1962). In a reversal of these magnitude estimation tasks, one could also give numbers to the assessor and ask him to find choice entities that match these numbers (e.g., find a choice entity that he would consider twice as valuable as a standard). Stevens (1975) calls these inverse methods magnitude production methods.

Bisymmetric measurement* -- Bisymmetric measurement formalizes the ideas represented in the procedure of bisection, described before, to a measurement theory formally justifying that method. The idea is to measure utility by bisecting intervals of choice entities (the word interval is used here rather loosely) into two equal parts, such that the utility differences between the bisection point and the two extremes are equal. Again, bisection theory is in principle applicable to all cases in Tables 2-4, but it can reasonably be applied only in simple cases, since the judgmental task involved in bisection may become very difficult if the choice entities are complex. We will first discuss the application of bisection theory to case I, and then sketch how the same ideas have been applied to case II by Fishburn (1975), who used suitable independence assumptions to simplify the bisection task.

The method of bisection itself defines the "operation surrogate"; the operation on two choice entities a and b is defined by finding an element c that bisects a and b. One wants, naturally, the property that c has the average utility of a and b in the numerical representation. The qualitative assumptions behind bisymmetric measurement are a little more complicated to spell out verbally than the ones for difference measurement. Again, as in difference structures, one wants the bisymmetry operation to behave nicely, for example, midpoints between a and b and between a' and b should preserve the preference order that existed between a and a'. (This is formally expressed as a monotonicity axiom in the Krantz et al., 1971, treatment of bisymmetric structures.) In utility measurement at least, one also wants the bisection

*Bisymmetric measurement has many different applications, among others, it applies to the measurement of utilities for two outcome gambles. The discussion of bisymmetric measurement here is restricted to the interpretation of the bisymmetry operation as bisection of utility intervals.

point of two choice entities a and b to be equal to that of b and a, and the midpoint between a and itself should be a. (These assumptions are called commutativity and idempotency.) Adding axioms that midpoints of midpoints behave nicely, too (the so-called bisymmetry assumption) and using possible associativity assumptions, one gets the following bisymmetric representation:

<u>Bisymmetric measurement</u>

(applied to bisection)

$a \succsim b$ if and only if

$u(a) \geq u(b)$ where

$u(a \circ b) = \frac{1}{2}u(a) + \frac{1}{2}u(b)$

where "o" stands for the bisection operation, and all other symbols have the usual meaning.

As mentioned before, the assessment procedure in bisymmetric measurement as discussed here would be of the bisection type described in the difference measurement sections as an approximation method. Also, all approximation methods discussed in that section should be good approximations for bisection measurement.

Fishburn (1975) applied bisymmetric measurement to cases more complex than case I. His motivation was to find appropriate assumptions that would guarantee that a bisymmetric utility function defined over these complex choice entities could be assessed as an aggregate of simpler bisymmetric functions defined only over some aspects of the choice entities. As an example, we will discuss here the bisection application of Fishburn's theory to the riskless multiattributed case II in Table 3.

Fishburn's models start exactly with the bisymmetric measurement model defined above. He then defines additional independence assumptions on preference orders and bisection operations in order for the bisymmetric function to decompose into single attribute functions. Fishburn's presentation of these assumptions is quite mathematical, but--in essence--they require that

1. some conditional preference orders are unaffected by the attribute values on which they are conditioned;
2. some conditional bisection operations are unaffected by the attribute values on which they are conditioned.

For example, if one would construct a utility function in one attribute using the bisection procedure, the shape of that function should not depend on the values at which the other attribute values were held fixed throughout that construction. These and similar assumptions produce the following four models, discussed -- in a slightly different form -- in Fishburn:

<u>Bisymmetric decomposition models</u>

$\underline{a} \succcurlyeq \underline{b}$

if and only if

$u(\underline{a}) \geqslant u(\underline{b})$

where (depending on independence assumptions)*:

1. Multilinear 1: $$u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i) + \sum_{i<j} c_{ij} f_i(a_i) f_j(a_j) + \sum_{i<j<k} c_{ijk} f_i(a_i) f_j(a_j) f_k(a_k) + \ldots + \prod_{i=1}^{n} c_{1,2\ldots n} f_i(a_i)$$

2. Multilinear 2: $$u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i) + \sum_{i<j} c_{ij} u_i(a_i) u_j(a_j) + \sum_{i<j<k} c_{ijk} u_k(a_i) u_j(a_j) u_k(a_k) + \ldots + \prod_{i=1}^{n} c_{1,2\ldots n} u_i(a_i)$$

---

*The model forms presented here generalize Fishburn's representations to n attributes. Fishburn's proof included only two attributes, but there are few theoretical difficulties in stepping to the n-dimensional case. Fishburn's proofs do not include the multilinear form (2), which is presented here because of its similarity to decomposable expected utility measurement, and because it could easily be derived in the bisymmetric context.

3. Multiplicative: $u(\underline{a}) = \prod_{i=1}^{n} u_i(a_i)$

4. Additive: $u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i)$

Here $\underline{a}$ and $\underline{b}$ are riskless multiattributed choice entities of the type described for case II. a_i, b_i are their respective values in attribute i. Note that the two multilinear forms include higher order interaction terms, which are either composed of the additive terms (2) or of independent terms (1). Practical assignment of utilities to choice entities within this framework proceeds as follows: first conditional bisection utility functions are constructed in each attribute using the bisection procedure described above. These functions are then interlocked (consistently scaled) by observing some additional indifferences between multiattributed choice entities, and aggregated according to one of the rules defined above, which depends on the independence assumptions postulated.

Conjoint measurement. -- Conjoint measurement theory as conceived in 1964 by Luce and Tukey and Krantz is probably the most prominent psychological measurement theory. So far its applications to utility theory are very limited, but it has a large number of potential application areas (conjoint measurement models can be found in 6 out of the 8 boxes in Tables 2-4). Conjoint measurement models are especially suitable for measuring utilities for choice entities that vary on several value relevant attributes, that have multiple affected individuals, or time variable consequences. Conjoint measurement has also been applied to choices among gambles as a special version of expected utility theory (Krantz and Luce, 1971, see also p.41). In the following, we will explain conjoint measurement via the example of measuring multiattribute riskless choice entities, but by appropriate substitutions for the word "attribute" (e.g., by "time periods", or by "individuals") the use of conjoint measurement for these other cases can be discovered.

Conjoint measurement constructs a utility function over multiattribute choice entities that decomposes into single attribute utility functions. The

type of decomposition and the rule by which these single attribute functions are aggregated depends on crucial independence assumptions in the model. So far, only "simple polynomial" combination rules to aggregate these single attribute functions have been axiomatized. The most prominent ones are the additive and the multiplicative rules. Other rules, not typically considered in decision analytic contexts, are distributive rules and dual distributive rules (see Krantz et al., 1971; Krantz and Tversky, 1971). Since the additive rule is by far the most attractive one for applied modeling purposes, (and since the multiplicative rule is -- in most cases -- a special case) the discussion of conjoint measurement will concentrate on this rule.

Conjoint measurement begins with a weak order defined over the set of choice entities. It then creates an "operation surrogate" by defining a choice entity c that expresses the combined effects of two other choice entities, a and b, together. This operation surrogate is the subjective equivalent of adding utilities.

The independence properties required to prove the additive conjoint measurement representation are usually called preferential independence. Preferential independence requires preferences over choice entities that vary only in some subsets of the attributes to be independent of constant values in the other attributes, no matter what the level of these constant values. Another way of saying this is that trade-offs in some subset of attributes are the same, no matter on what constant values in the remaining attributes these trade-offs are conditioned. Yet another way of stating this requirement is by referring to the actual construction procedure. Utility function constructed while values in some attributes are held fixed should have a shape that is independent of that fixed value. In particular, any of the single attribute utility functions should not depend on these conditional values. For example, in evaluating sites for a nuclear power plant, the utility cost function over the attribute "population density in a twenty mile radius" is probably independent of, say, "cost of transmission lines" for that particular site. Transmission lines costs and costs for access transportation are probably jointly preferentially independent of population density, etc. For some counterexamples, see v. Winterfeldt and Fischer (1975).

Conjoint measurement (simple polynomials)

$$a \succsim b$$
if and only if
$$u(a) \geq u(b)$$

where (depending on independence assumptions):

1. Multiplicative model: $u(\underline{a}) = u_1(a_1) \cdot u_2(a_2) \cdot u_3(a_3)$

2. Distributive model: $u(\underline{a}) = u_1(a_1) \cdot u_2(a_2) \cdot + u_3(a_3)$

3. Dual distributive model: $u(\underline{a}) = u_1(a_1)\,[\,u_2(a_2) + u_3(a_3)\,]$. Here, $\underline{a} = (a_1, a_2, a_3)$.

Construction of the recurring utility functions u_i in these conjoint measurement models is somewhat similar to the standard sequence approach in difference measurement. This procedure has occasionally been called "dual standard sequence procedure" (Krantz, 1964), or "saw tooth procedure" (Fishburn, 1967), or "lock and step procedure" (Keeney and Raiffa, 1975). It uses indifference judgment and constructs matches between choice entities that vary only on two attributes (events, time-periods, individuals) at a time. A unit step in attribute 1 is used to lay off a sequence of steps in the other attribute with the first attribute held at a fixed level. This insures that the elements in that sequence space out the attribute in equal utility steps.

For example, when evaluating an apartment on the two attributes, rent and driving distance, one considers questions like the following: how much more driving time are you willing to add to, say, 5 minutes, for a rent decrease of from \$210 to \$200? If the answer is 10 minutes, one continues by asking: how much driving time are you willing to add to 15 minutes for a rent decrease of from \$210 to \$200? If the answer is 13, one can infer that the increase in utility cost from 15 minutes to 28 minutes is equal to the increase from 5 to 15 minutes driving time. This way a good number of points of the utility function can be assessed and a curve can be smoothed through these points. Theoretically, the spacing of the points in the single attribute can be made arbitrarily fine in order to increase the goodness of approximation of the whole utility function. One would just use a smaller unit in

the comparison attribute. Practically, however, if one decreases this unit, certain discrimination problems may arise. For example, if one would try to lay off a standard sequence over the attribute "driving time" against a unit step in the attribute "rent" from $201 to $200, the assessor may be hard pressed to come up with reasonable answers.

After the construction of the single attribute utility functions, one still has to make sure that the different functions u_i have comparable units. This can be done by determining indifferent choice entities whose indifference is not already implied by the previous construction, and by solving the resulting equalities.

The construction procedure of dual standard sequences resembles that of bisection and that of difference sequences, and similar judgmental processes are required to make the indifference judgments which create these sequences. One can expect closely resembling shapes of the utility functions using any of the three procedures for assessing utility functions in a conjoint measurement context, although, theoretically speaking, only dual standard sequences are appropriate. A special problem arises when one uses bisection or difference judgments, or any of their approximations (rating scales, method of equal appearing intervals, etc.) to approximate single attribute utility functions in the conjoint measurement context. These functions are constructed disregarding any other attributes, and therefore will have to be carefully matched in their units to ensure comparability. One could use the formally "clean" solution of observing a number of independent complex indifferences, and solve the resulting equalities, just as in constructing utility functions with standard sequences. An alternative is to assess these scaling factors or weights directly by a magnitude estimation procedures (direct rating of weights between 0 and 100; distributing 100 importance weights points among the attributes, etc.). This procedure has been used by Huber et al. (1971), and v. Winterfeldt and Edwards (1973). Edwards (1971) describes a special version of magnitude estimation for importance weights in which the decision maker assesses the ratios of the weights for two attributes at a time. From these, all weights can be inferred. Since scaling factors have the property of a ratio scale, this appears to be a reasonable procedure. The problem

with such explicit numerical weighting schemes is that they do not directly focus on comparisons of utility differences and thus may depend on the relative ranges of the values in the single attributes. But this problem can be avoided through a careful assessment of these relative ranges and/or by making the assessor aware of that range (see Keeney and Raiffa, 1975, for a discussion of this point in a slightly different context).

Expected utility theory -- Expected utility theory could be called the theoretical cornerstone of decision analysis. Although some may consider decision analysis simply an application of expected utility theory, the previous sections should have made clear that there are many concepts of utility which could be applied in a formal analysis of decision problems.

Expected utility theory has been applied to all risky cases in Tables 2-4. Although there are a large number of expected utility axiomatizations, there appear to be only four basically different approaches to measure the utility of risky choice entities with expectation models.*

These are:

1a. V. Neumann and Morgenstern's (1947) expected utility theory with numerical probabilities;

1b. Savage's (1954) subjective expected utility theory;

1c. Davidson, Suppes, and Siegel's (1957) finite utility theory for equally likely events;

1d. Luce and Krantz's (1971) conditional expected utility theory.

Let us, however, first state the similarities among the four approaches and then discuss the differences. First, EU-theories all make--in one version or another--three crucial assumptions about preferences among risky choice entities. The first is the common weak order assumption, which was discussed recurrently in previous sections. It says that the decision maker can order risky alternatives transitively. The second belongs to the class of independence assumptions mentioned earlier. It takes different forms

---

* A fifth possibility, not included here is an application of bisymmetric measurement to risky options in which the bisymmetry operation a ∅ b would be interpreted as "receive a if one event occurs, b otherwise". (see Pfanzagl, 1968).

in the four approaches, but it is commonly known as the sure thing axiom. The sure thing axiom requires that preferences among risky alternatives should be independent of events in which these alternatives have common outcomes. This assumption justifies the additive form of the expected utility model, just as preferential independence justified the additive form in conjoint measurement. The final assumption belongs to the class of "technical" axioms, and it is a combination of both a solvability condition and an archimedean condition. It requires that no choice entity is infinitely desirable or undesirable, and that there are certainty equivalents for all possible uncertain entities (i.e., that the decision maker is able to find a riskless entity that is just as valuable to him as the risky entity).

If these conditions are met, a utility function over risky options can be constructed that has the following properties:

<u>Expected utility measurement</u>

$$\underline{a} \succcurlyeq \underline{b}$$

if and only if

$$u(\underline{a}) \geqslant u(\underline{b})$$

$$\text{where } u(\underline{a}) = \sum_{j=1}^{n} p(E_j)u(a_j)$$

Here $\underline{a}$ is a risky choice entity (a gamble) whose outcome is a_j (e.g., a certain \$-amount) if event E_j occurs. $p(E_j)$ is the numerical probability of event E_j.

All four approaches end up with some measurement representation similar to the one above. The specific form and interpretation, as well as the theoretically feasible construction methods differ, however.

v. Neumann and Morgenstern's theory is the classic expected utility theory. Its main restriction is the assumption that numerical probabilities are known for all events. These numerical probabilities play, in some sense, the role of an operation surrogate in their theory. "apb" (i.e., the p-operation "put a and b together") would be interpreted as "with probability p you will receive a, with probability 1-p you will receive b". Since probabilities are assumed to be known, utilities can be constructed within the v. Neumann and Morgenstern framework by observing indifferences between lotteries and sure

outcomes. If a, b, c are riskless outcomes, and (apc) is the gamble that yields a with probability p and c with probability 1-p, and if b is indifferent to (apc), then p is a measure of the utility of b relative to the utilities of a and c. By arbitrarily assigning utility values of 0 and 1 to two choice entities, such indifferences imply equations through the expected utility representation that can be solved for the unknown utilities. For example, if the utilities in the above case were 1 for a and 0 for c, then the expected utility representation would imply that the utility of b is p.

Indifferences can be observed either by varying the probability p in (apc) and holding b fixed; or by fixing p and varying b. If the choice entities have some numerical description (such as units of a commodity), it is often sufficient to determine the utilities for only a few points and approximate the utilities for intermediate points by interpolation. This general type of utility construction through indifference lotteries with known probabilities is probably the most common procedure in decision analysis--although, as this report demonstrates, it is by far not the only one.

The main problem with v. Neumann and Morgenstern's expected utility measurement is the assumption that probabilities of events are known. Savage overcame that problem in an ingenious way. In essence, he combined earlier theories of subjective probability measurement (Koopman, 1940) and v. Neumann and Morgenstern's theory. All his assumptions are expressed in form of preferences among uncertain alternatives which are described as a set of outcomes a_i to be received conditional on the occurrence of a particular event E_i. No numerical probability is assumed for these events. Using these preferences, Savage constructed an induced relation among events, which is interpreted as the relation "more likely". Then he made use of the fact that the probability of events can be measured not unlike length in an extensive measurement model (see Koopman, 1940). Events can be compared (with the "more likely" relation) just as rods can (by the "longer" relation). An operation also can be defined for events, namely the union of two mutually exclusive events (just like two rods can be connected). If the independence assumption holds, that the relation among two events is preserved when both are united with a third event, then one can show that a numerical probability representation

exists. Savage's ingenious idea was to express these independence assumptions for the induced "more likely" relation in form of preference relations, which allowed him to construct a numerical probability for events. He then proceeded to make use of this numerical probability in proving the expected utility theorem in essentially the same way as v. Neumann and Morgenstern did.

To construct utilities in this context, one therefore has to first construct probabilities, and then use the v. Neumann and Morgenstern procedure described above to construct utilities. The formally justified construction of probabilities for events is again that of a standard sequence, that compares the union of many equally likely events with the event to be measured. If the union of n of these equally likely events is "equally likely" as a certain event, and the union of m of these events is "equally likely" as the unknown event, then the unknown event has a subjective probability of $\frac{m}{n}$.

Davidson, Suppes, and Siegel (1957) went yet another route in modeling expected utility theory. Their theory resembles closely the difference measurement that has been described before. In essence they built a difference structure using equally likely events by defining the "difference" between choice entities a and c to be equal to that between b and d if a gamble that yields with equal likelihood a or b is indifferent to a gamble that yields with equal likelihood c or d. (The meaning of this definition can easily be inferred by using the expected utility representation.) Then they formulated axioms on the preferences among gambles with equally likely events that allow identification of a difference structure. Construction of the utility function resembles that of difference measurement: a sequence of indifferent gambles with equally likely outcomes is created that lay off a sequence of outcomes with equal utility differences. Note, however, that utility "difference" has a different meaning here than in the direct difference measurement.

Luce and Krantz used yet another measurement framework to construct an expected utility theory. They applied conjoint measurement theory to evaluate the utility of risky choice entities. Their motivation was to get around a property of Savage's model which lies in the description of the choice entities as acts that produce different consequences given the <u>same</u> set of events. This view of the choice entity is best characterized by the usual

description of the decision problem under uncertainty as a payoff matrix, a matrix in which each row is an act, each column is an event, and each matrix entry is an outcome to be received if an act row is selected and an event column occurs. This representation implies that no matter what act is selected, the decision maker will always face the same uncertainties. Luce and Krantz point out that few problems are of this kind (although most problems can be defined into that format), and that it is often more convenient and sometimes unavoidable to formulate the problem in terms of conditional acts, i.e., decisions that are conditional in the sense that by choosing them, one is restricted to some subset of all events bearing on the decision problem. Luce and Krantz then basically create a conjoint measurement system in which they measure the utility of these conditional acts given the restriction on the sets of events. Besides the usual preferential independence assumptions which only guarantee the additivity part of the resulting expectation model, they make assumptions that guarantee that the conditional utility functions are differing only in units. This assumption is spelled out in Luce and Krantz, and basically requires a standard sequence laid off conditional on one event to be a standard sequence in any other event too.

The formally justified construction of the utility of conditional decisions follows the pattern sketched in the discussion of conjoint measurement. Standard sequences are built conditional on each event to construct the conditional utility functions. Probabilities are inferred from the comparison of the units of these standard sequences.

In practice, utility functions are, of course, not constructed by using difference equations, standard sequences of equally spaced outcomes, or standard sequences in conditional events. The most common procedure is to numerically estimate probabilities or probability distributions of risky alternatives, and then to use indifference lottery procedures to construct utilities. These construction procedures have changed little, if any, from the theoretical framework v. Neumann and Morgenstern suggested, although the more recent models ask for quite different procedures. Strictly speaking, these separate assessments of numerical probability and utility are approximation methods to construct utility functions in the framework of Savage, or Luce and Krantz.

We turn now our attention to the cases VI, VII and VIII, in which various special forms of expected utility theory have been applied. There are two ways in which expected utility theory can be used to measure utility in these cases. Both assume an expected utility representation over the basic choice entities (risky and multiattributed, risky and time variable, or risky and multiple individuals affected). The first approach uses some riskless assumptions such as weak order, convexity of indifference curves, preferential independence, etc., to construct a riskless utility function. (The literature often calls such a function "value function" or "ordinal utility function".) Then a function h is constructed that transforms the riskless function u into a risky utility function u' that follows the expected utility principle. These are the models 1 and 2 in the multiattribute and the time variable cases VI and VII. In the multiattribute case one may also consider a riskless bisymmetric utility function that is transformed into a risky function. (Models 3a-d in case VI). None of these types of models has been considered yet for the case VIII in which many individuals are affected by a decision.

As an example of this approach, consider the model 2a in case VI, the additive conjoint measurement expected utility model. If preferential independence is satisfied, a riskless utility function u can be assessed that is of the form

$$u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i)$$

where the a_i's are values of $\underline{a}$ in the attribute i. The construction of this utility function would use dual standard sequences or appropriate approximation techniques. Then risky utilities for various points of that utility function can be assessed using standard lottery procedures to generate the transform h that gives

$$u'(\underline{a}) = h(u(\underline{a})).$$

If the $\underline{a}$'s themselves are uncertain, that is if one receives the multiattributed outcome $\underline{a}_j$ conditional on the occurrence of an uncertain event E_j, the expected value of u':

$$E = \sum_{j=1}^{m} p(E_j)\cdot h(u'(\underline{a}_j)) = \sum_{j=1}^{m} p(E_j)\cdot h\left[\sum_{i=1}^{n} u_i(a_{ij})\right]$$

preserves the order of preferences among these uncertain and multiattributed choice entities.

The second approach to modeling risky multiattributed (time variable, multiple individuals) preferences is labelled "expected utility decomposition models". These models make independence assumptions about preferences among risky choice entities beyond just the simple EU-model. Given these assumptions the risky utility function of complex choice entities (risky consumption streams, risky multiattributed objects, risky decisions that affect many) can be decomposed into utility functions of single attributes, time periods or individuals.

These single attribute utility functions are then aggregated according to some rule that depends on the types of independence assumptions made. Since these rules are formally identical for cases VI, VII, and VIII, only the multiattribute case will be discussed here in detail. The four most common model forms are very similar to the bisymmetric decomposition models (but note that in bisymmetric measurement nothing guarantees that the resulting utility function will be appropriate for taking expectations):

4a. Multilinear 1: $u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i) + \sum_{i<j} c_{ij} f_i(a_i) f_j(a_j) + \ldots$

$$+ \prod_{i=1}^{n} c_{1,2,\ldots n} f_i(a_i)$$

4b. Multilinear 2: $u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i) + \sum_{i<j} c_{ij} u_i(a_i) u_j(a_j) + \ldots$

$$+ \prod_{i=1}^{n} c_{1,2,\ldots n} u_i(a_i)$$

4c. Multiplicative: $1 + ku(\underline{a}) = \prod_{i=1}^{n} (1 + ku_i(a_i))$

4d. Additive: $u(\underline{a}) = \sum_{i=1}^{n} u_i(a_i)$

Here, $\underline{a}$ is a (riskless) multiattributed object, a_i is its value in attribute i, the u_i's and f_i's are single attribute utility functions, and the c_{ij}'s, k, etc. are scaling factors. The expected value of u preserves the order of preferences, if the $\underline{a}$'s themselves are uncertain choice entities.

The strongest of these models is the additive expected utility model (4d) first conceived by Fishburn (1965) and Pollak (1967). It requires that preferences over risky choice entities that differ only in some subset of the attributes should be independent of constant values or constant lotteries in the remaining attributes. Another way of saying the same thing is that preferences among risky choice entities ought to depend only on the marginal probability distributions in single attributes. This assumption has also been called the "marginality assumption", "additive independence" or "value independence".

The multiplicative model was developed by Pollak (1967) and by Keeney (1968, 1974). It weakens the marginality assumption by requiring that preferences over risky choice entities that differ in some subset of the attributes should be independent only of constant values in the remaining attributes. That is, they may depend on lotteries in the remaining attributes. This axiom is usually called "utility independence". (See also Fishburn and Keeney, 1974, for several weaker versions of this assumption.)

An even weaker assumption leads to the multilinear model 4b, developed by Keeney (1968). The multilinear model makes assumptions only about conditional preferences in single attributes. Preferences over risky choice entities that vary only in one attribute should be independent of constant values in the remaining attributes (see also Farquhar, 1974). This single attribute version of general utility independence has not been specifically named in the literature.

The multilinear model 4a generalizes Keeney's multilinear model by allowing independent interaction terms. This model has been developed by Fishburn (1973, 1974) using assumptions similar to those in the development of the multilinear bisymmetric model. These assumptions are quite difficult to spell out,

and their intuitive meaning is by no means obvious. For a discussion of this and some even more complex models, see Fishburn (1974), and Farquhar (1974 a and b).

Going in the opposite modeling direction, some specifications of the additive and multiplicative expected utility models are discussed in Meyer (1969) for the risky time variable case VII. Meyer's assumptions that preferences are not only utility independent, but also time stationary imply that utility functions u_i (the single time period utility functions) vary only in unit, not in shape. An additional assumption guarantees that the units decrease exponentially over time. The results are the multiplicative and additive models with variable or constant discounting rates (3a and 3b in VII).

The formally justified assessment procedures to construct any of the models 4a-4d are all based on the type of indifference lottery procedures discussed earlier. These assessments are done in single attributes while the remaining attribute values are held fixed. Scaling constants c_{ij}, k, etc. and techniques to match the utility functions u_i in units require indifference lotteries that involve more complex choice entities, varying on at least two attributes simultaneously. Depending on the acceptable independence assumptions, any of the model forms 4a-4d is then applied to aggregate the single attribute utility functions. These construction procedures are best described in Keeney and Raiffa (1975) for models 4b-4d, and in Fishburn (1973c, 1974a) for model 4a.

The literature is rather silent about possible techniques to approximate decomposable expected utility functions. One possibility is to approximate the risky utility function by a riskless utility function assessed with traditional riskless methods and then compute the expected value of this riskless function to determine an approximate utility for risky choice entities. If it is assumed that a riskless utility function is a reasonable approximation to the risky function, judgmental methods such as rating scales can be used to approximate the riskless function. Another way of dealing with the problem of approximating risky utilities with riskless utilities is to make the risky problem a riskless one by determining for any risky choice entity a riskless one

that is indifferent to it. Such certainty equivalents could be assessed, for example, in each attribute and then the utility for the now riskless choice entity could be measured with appropriate riskless procedures. (See also Keeney, 1968, and Keeney and Raiffa, 1975, for a discussion of model assumptions that allow conversions of risky objects into riskless ones by assessing certainty equivalents in single attributes.)

Relationships between models and assessment procedures

This part of the report will show some additional relationships that link the 5 model classes and show connections between models for different choice entities. In order to avoid unnecessary duplication, only the riskless and risky, single attributed and multiattributed applications of these models will be discussed in detail. Other applications to group and time problems can be inferred from the single attributed and multiattributed distinction by substituting "individuals" or "time periods" for "attributes".

We are interested here in conditions under which models and assessment procedures produce the same or possibly very similar utility functions. There are at least three arguments leading to such coinciding or converging utility functions.

The first examines the logical relationships between models. Two models A and B are equivalent if the assumptions of model A imply the assumptions of model B and vice versa. Utility functions constructed within equivalent models should be indistinguishable. In a weaker relationship, model A may imply B, but the reverse may not be true. If in such a case the assumptions of model A are true, then by implication also the assumptions of model B will be true. Utility functions constructed within either model should coincide in the sense that they both have the properties required by the weaker model. An example of such a case is the additive expected utility which implies the multilinear expected utility model(s), but the multilinear models do not imply the additive model. Another example is additive bisymmetric measurement which implies conjoint measurement but not vice versa.

The second argument for agreement between utility functions constructed with different procedures and models is based on the similarity of the judgmental processes involved in construction procedures. For example, although there is no logical reason to assume that single attribute difference measurement and single attribute bisection measurement should produce identical utility functions, one would not expect them to differ very much since the processes involved in bisection are not substantially different from those involved in making difference judgments. This line of reasoning can even conclude in similarities between risky and riskless utility functions.

The third argument is less subtle, requiring neither logical nor behavioral similarities. It is based on the experience (either experimental or through simulation) that certain models and procedures will produce converging utility functions in a large number of cases. For example, it has been shown that additive models are usually pretty good approximations of nonadditive models; and that variations in single attribute utility functions or weighting parameters produce utility functions that are very highly correlated. This type of insensitivity is often used to justify a model whose assumptions are not met or not checked, or to apply procedures that are formally not justified.

Formal model implications and equivalences -- Models for the risky and riskless single attribute case have few interesting formal relationships. In the riskless case, clearly difference measurement and bisymmetric measurement imply the weaker weak order model, but not each other. In the risky case, the only interesting relationship links weak order risk theories and expected utility models. As Coombs (1972) points out, expected utility models imply the weak order expected risk model (portfolio theory). In other words, if the expected utility model fails, it may still be possible to assess preferences with a risk model.

Tables 7 and 8 summarize all implications for models for risky and riskless multiattribute cases. Identical charts are applicable to the other two cases (individual vs. group; time invariant vs. time variable). The direction of arrows in these charts goes from the stronger model to the implied (weaker)

model. Some of these implications are rather obvious and can be directly inferred from the model forms, e.g., the fact that the multilinear model 4a with independent interaction terms is implied by the multilinear model 4b with dependent interaction terms, or that the additive bisymmetric model implies the multilinear forms.

Insert Tables 7 and 8 about here

An interesting case is the implication of the additive and multiplicative bisymmetric models to the additive conjoint measurement model. Both bisymmetric models are stronger than the additive conjoint measurement model, since they make assumptions beyond just the preference order of multiattributed choice entities. But their assumptions (together with appropriate continuity assumptions) imply that there is an additive order preserving representation for multiattributed alternatives which implies conjoint measurement. (In the multiplicative bisymmetric case, this additive representation would be a logarithmic transformation of the bisymmetric function).

The risky multiattribute case has some even more intriguing model implications to offer. The strongest model here is the additive expected utility model that implies almost every other model (with the exception of some simple polynomial conjoint measurement decompositions). Since the bisymmetric decomposition models have identical form and differ from the expected utility decomposition models only in the transformation h that maps their utility functions into a risky utility function, all of these models are implied by the respective expected utility decomposition models. Most of these relationships follow directly from the functional form of the models.

What about the relationship between risky and riskless models? For the single attribute case the only obvious implication is that any model in the risky case generates utility functions that preserve also the order of riskless preferences, that is, they all imply the weak order model in case I. However, nothing guarantees that EU-models imply difference or bisymmetric models (although an argument will be made later that there are certain behavioral similarities between these models).

Table 7: Implication chart of models for riskless and multiattributed choice entities

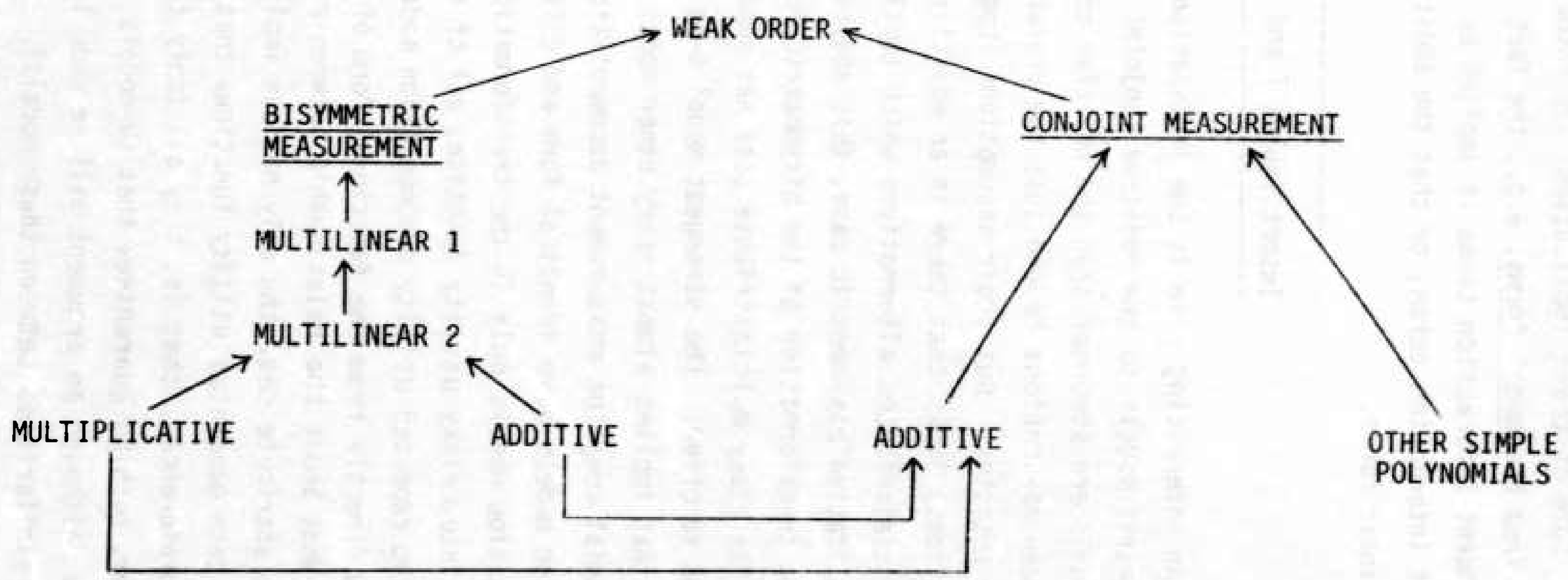
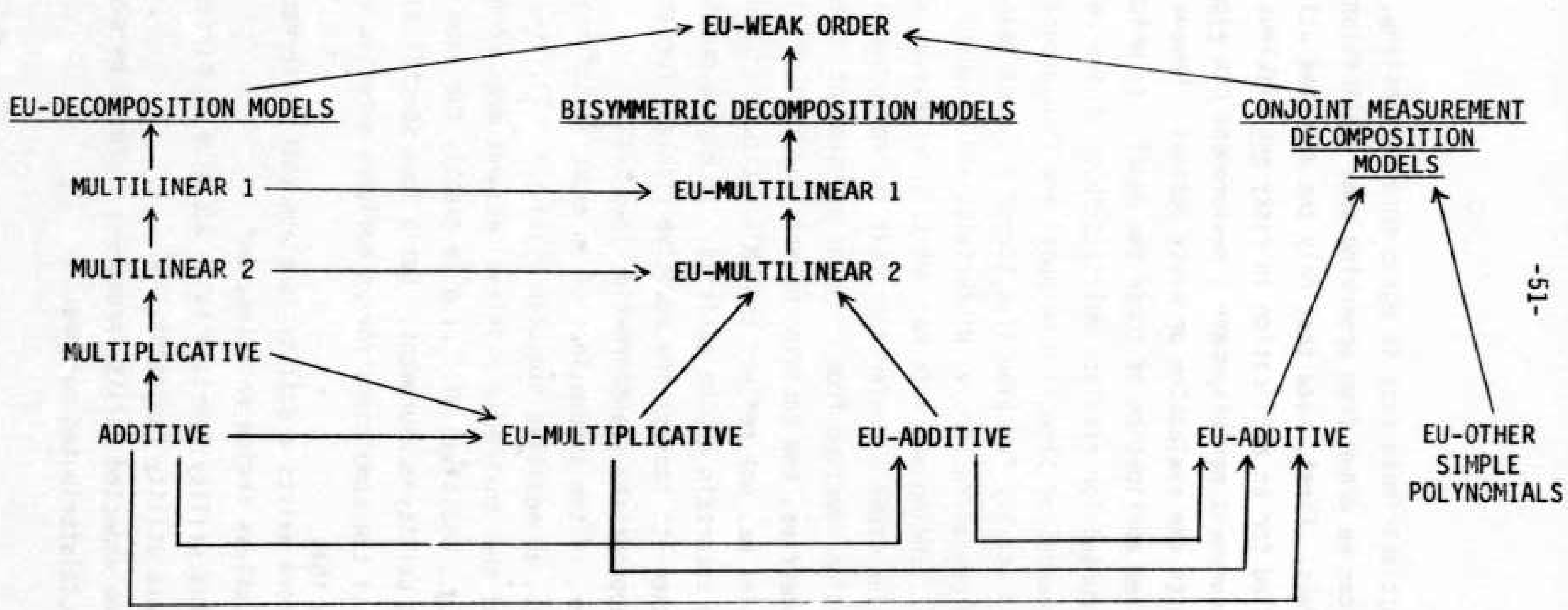


Table 8: Implication chart of models for risky and multiattributed choice entities



The multiattribute case is again more interesting, and some practical conclusions can be drawn from observing some implications between risky and riskless models. First, note that only the expected utility models can be formally justified for an application in risky <u>and</u> riskless situations. Neither conjoint measurement nor bisymmetric measurement can <u>alone</u> provide a mechanism for coping with the evaluation of risky options. However, one could conceive of a simplified application of these two models, in which first a utility function is contructed for riskless multiattribute choice entities (using either standard sequences or bisection methods) and then expectations are taken of that riskless utility function if outcomes are uncertain. This is <u>not</u> in general a valid procedure (see v. Winterfeldt and Fischer, 1975). To give an intuitive understanding why this may not be appropriate, consider the properties of the decision maker's preferences that are reflected in an additive riskless utility function, derived from conjoint measurement. These are riskless independence properties, the decision maker's marginal utilities, etc. However, these utilities may not reflect the DM's attitude towards risk when he has to choose among uncertain choice entities. It may be necessary--before expectations are taken--to convert the additive riskless function into a risky function by an appropriate transformation (see p.43).

However, if the assumptions of an expected utility decomposition model are met (e.g., an additive expected utility model), then a formal argument can be made, that the analogous riskless conjoint measurement or bisymmetric measurement (e.g., additive) will produce exactly the same utility function as the expected utility measurement. Let's look specifically at the additive case here. If the conditions for an additive expected utility model are met, then we know that

1. there exists a utility function that preserves preferences over riskless choice entities, and
2. this utility function is an additive combination of single attribute utility functions, and
3. the expected utility preserves preferences over lotteries with multiattributed outcomes.

Now assume that, in addition to the axioms of additive expected utility theory, also the axioms of conjoint measurement and the axioms of an additive bisymmetric structure hold. These assumptions imply two more utility functions that

1. preserve preferences over riskless choice entities, and
2. decompose additively over attributes.

The next step is a consequence of the uniqueness property of all three utility functions, which says that any two utility functions that preserve preferences over riskless alternatives and decompose additively, must be linearly related to one another. Thus all three utility functions must be linearly related. The real trick comes now: if, however, a conjoint measurement utility function, a bisymmetric utility function, and an expected utility function are linearly related, their respective expectations must preserve preferences over lotteries with multiattributed choice entities. Practically, speaking, this means that given the conditions of additive expected utility theory, <u>any</u> model that produces preference preserving additive utility functions could be used as a surrogate for the original expected utility model, and its utility functions could be used to take expected values. So, if an analyst has convinced himself of the validity of the additive EU-model, he might as well go ahead and construct utilities with standard sequences or bisection methods. (And, of course, the step from here to even simpler approximation methods is not very big.)

A final example of this sort is the case in which the multiplicative expected utility model is valid. In this case, the additive conjoint measurement model for riskless choice entities will be valid, and both utility functions will be related by a logarithmic (or exponential) transformation. So, it is perfectly valid -- when a multiplicative EU- model is accepted -- to construct utilities with conjoint measurement procedures and to transform these utilities exponentially to achieve the multiplicative EU representation.

<u>Behavioral similarities and differences in assessment</u> -- Although quite a few models that were discussed so far cannot be logically related, it is possible to look at the cognitive processes that are involved in the assess-

ment task to see if there are any reasons to assume that utility functions resulting from different procedures may produce similar results. This point is particularly important for the justification of approximate assessment procedures, which are often much simpler than the theoretically feasible ones. Unfortunately, the experimental and applied literature is rather silent about such similarities, so most of what follows will be speculative.

Let us first look at models and procedures in the single attribute case. In the riskless case we already discussed some aspects of behavioral similarities of models and procedures. Standard sequences of utility differences, direct rating scales, bisection procedures, methods of equal appearing intervals --all the theoretically feasible and approximation methods for this case--appear to involve similar judgmental processes: that of judging relative preference differences among riskless choice entities. Although cognitive processes such as anchoring and adjusting or context effects may operate differently in these procedures, it would be surprising if large differences in the shapes of the single attribute utility functions were to be found.

Among the models for the risky single attributed case (III), the expected utility models are all constructed using some versions of indifference lottery procedures. Whether one varies probability distributions in these procedures and asks for certainty equivalents, or whether one varies outcomes and asks for matching probabilities is irrelevant from a theoretical point of view. Practically, these two types of indifference lotteries may produce different results, partly because of their shift in emphasis on different aspects of the lotteries (costs and payoffs vs. probabilities). The most interesting question arises in comparing the behavioral similarities and differences among procedures that jointly measure probability and utility and procedures in which the assessment of probability and utility are strictly separated. None of the three SEU-models would theoretically justify numerical estimation of probabilities and subsequent assessment of utilities with the use of these numerical probabilities. Nevertheless, this is the most widely used approximation procedure. Magnitude estimation of likelihoods of events in the form of odds or probabilities, although not formally justified by the models, in-

volves, of course, processes that are very similar to the processes in Savage's standard sequence method. All one has to assume is that the assessor has some event in mind that is well calibrated (such as the event that a spinner will land on a certain section of disc), finely divided in many equally likely sub-events, and that his numerical assignment is an indifference judgment between such well calibrated events and the event in question. There are, of course, many other procedures to assess probability distributions for risky options, but since this paper is concerned with utility rather than with probability, it will not go into further detail.

How can models and assessment procedures for the riskless and risky single attribute cases I and III be compared behaviorally? Here, I would like to express a rather heretical view. Certain cognitive processes involved in the judgments of utility differences resemble closely those involved in judging gambles. For example, a decision maker may reject playing a fifty-fifty gamble for $100, because he judges the loss of $100 as more severe than he appreciates winning $100. The risk aversion he exhibits by not playing the gamble really is based on judgments of utility differences. Very often risk attitude (or the shape of a risky utility function) can be explained by the characteristics of a riskless utility function, e.g., its marginal utility, rather than by a special component introduced through gambling. For example, if a decision maker has the option to play a gamble in which he receives 10 pounds of ground beef on the flip of a coin vs. nothing, he may state a certainty equivalent for that gamble of 3 pounds -- not because he is risk averse (in the non technical sense that he does not want to take chances) but because he does not see any value in an additional pound of ground beef beyond, say, 6 pounds.

In indifference judgments about gambles, riskless judgments of utility differences and pure risk attitude are confounded. The above examples suggest that riskless utility differences may play a very strong part in the judgment of risky utilities. If pure risk attitude is defined through the transformation of a riskless difference utility function into a risky one, that transformation may often be almost linear. In the multiattribute risky and riskless situation, this linearity can even be proven provided that an additive

expected utility decomposition model holds (See pp.52-53). In summary, local risk aversion and marginal utility may coincide more often than the conceptual distinction of the underlying models may suggest.

Bisection procedures and indifference lottery procedures can be similarly linked. If a certainty equivalent of a gamble with two equally likely outcomes is identical to the bisection point for all possible outcome combinations, then a bisymmetric utility function and a risky utility function would have equal shape. Again, it appears that the processes leading to the identification of a bisection point are not that much different from the processes leading to a certainty equivalent.

From these arguments, the next step -- as radical as it would first appear -- seems not that far: rating scales involve similar cognitive processes as do bisection and difference judgments. Bisection and difference judgments may involve similar processes as do indifference judgments about gambles with two outcomes and equally likely events. So there are behavioral reasons to assume that rating scales will produce utility functions that are not substantially different from utility functions generated with indifference lotteries.

All these concepts which have only been sketched above could, of course, be axiomatized, and the conclusion could be proven, and experimentally checked. So far, it is only speculation intended to break up some of the rigid thoughts about what utility means and how it should be measured. Later arguments about the insensitivity of certain models will show, anyway, that precise utility and probability assessment probably does not matter very much.

The multiattribute case really adds only one new assessment procedure: that of dual standard sequences. All other procedures are single attribute versions of the already discussed methods (bisection, and indifference lotteries). It is doubtful that standard sequences should involve cognitive processes that are very different from, say, difference judgments. Really, what the assessor eventually does in creating a standard sequence is to find choice entities that are equally spaced in utility. The only difference from the difference standard sequence procedure is that he does so by using a standard comparison step <u>in another attribute</u>. One could argue that dual standard sequence procedures are difference procedures with the help of a realistic representa-

tion of a utility difference. The argument can then be made that both bisection and direct rating procedure will generate results similar to dual standard sequences.

One crucial difference which arises only in the additive multiattribute case is that of implicit vs. explicit weighting procedures. Recall that implicit weights are calculated from equations that result from observed indifferences among rather complex stimuli. The alternative approximation method is direct rating or ratio assessment of such weights. Here, it is possible that different cognitive processes are operating when making such judgments. When a decision maker has to make indifference judgments which eventually allow the computation of weights, he will express his local trade-off between attributes. This local trade-off (which is, of course, variable with the location of the two choice entities that are matched) allows the identification of the trade-off in utility as measured on the unrescaled single attribute utility functions. This trade-off in utility tells how many (unrescaled) utility units the decision maker is willing to give up in attribute 1 for an increase of x (unrescaled) utility units in attribute 2. Since this trade-off is constant in utility, that is enough information to get the utility units of the two unrescaled utility functions into correct proportion. So really, the processes that are tapped here when observing indifferences to construct rescaled utilities are directly related to comparisons of utility intervals.

When importance weights are judged directly, however, either on a numerical rating scale, or in terms of ratio magnitude estimation, factors other than comparison of utility units may enter into the decision maker's consideration. One possibility is that the range within which the single attribute utility functions are assessed is disregarded--which in essence is disregarding the size of unrescaled utility intervals in that attribute--when judging some absolute "importance ratio". In attributes that have an insignificant range, this will lead to overestimation of the rescaling factors; in attributes that have a wide range, this may lead to underestimation. In any case, external factors not related to the scaling problem may enter in judgments of importance ratios.

<u>Similarity by insensitivity</u> -- The final argument which supports convergence among several models and procedures is that of insensitivity, a sort of de facto similarity without formal or behavioral cause. It just says: experience has shown that model A and model B or that procedure a and b produce converging utility functions. Most of these results have been developed for additive models, but there are also some indications of model convergence across the borderline of additivity. Fischer (1972) and Yntema and Torgerson (1961), for example, demonstrate that additive models can approximate non-additive models quite well. Similar arguments can be found in Dawes and Corrigan (1973), who introduce the qualification "if the dependent variable (utility of the non-additive model) is measured with a substantial amount of error". There is a wealth of regression analytic literature showing that simple linear models produce surprisingly good results when compared with more "realistic" figural and complicated models. Fischer (1972), however, found some examples, where additive models are not such good approximations, in particular for complex multilinear models when the number of attributes become large.

Most recent insensitivity research was concerned with convergence between additive models with different utility functions or weighting parameters. The main results of these studies are:

1. Variations of the shape of single attribute utility functions will produce overall utilities that are highly correlated as long as all single attribute functions are monotone* (Fischer, 1972); (see also Slovic and Lichtenstein, 1971, for similar arguments in regression analysis).
2. Variations in weight parameters produce overall utilities that are highly correlated. Unit weighting schemes often do a remarkable job in predicting models with skewed weighting schemes (Einhorn and Hogarth, 1974; Dawes and Corrigan, 1973.

---

*Monotonicity is a version of preferential independence that requires that more of <u>one</u> attribute is always preferred to less (or vice versa), no matter what the other attribute values are.

Although few of these analyses have concentrated specifically on utility models (most of them were concerned with simple regression models, in which one problem was the consumption of degrees of freedom in parameter estimation, a problem that does not arise in utility theory), the results should have implications for assessment in utility theory. It is probably not too daring to say that, as long as all single attribute functions are monotone, their precise shape and relative rescaling will not matter very much.

How useful is utility theory for decision analysis?

So far, we have classified, described, and integrated the formal models and assessment procedures that form under the name of utility theory. This section examines the question: what is all this good for? Is the effort that measurement theorists put into the development and sophistication of utility models really useful to anybody, or is it just a mathematical exercise? Obviously, Hölder did not produce any changes in the practice of length measurement. Do utility theorists run a similar danger of reproducing results that really have only little, if any, implications for measurement practice?

To answer this question, let us first look into some areas where utility theory has been applied to model preferences. Much theorizing and psychological experimentation has gone into the analysis of utility theories as descriptive models of human decision making behavior. Studies in the descriptive validity of utility theories have concentrated on two decision situations: the riskless multiattributed case (case II in Table 4) and the risky single-attributed case (case III in Table 4). Simple additive models have been tested in the riskless multiattribute case; they usually were considered approximations of some conjoint measurement representation (see Fischer, 1975). Expected utility models and conjoint measurement models have been applied to the risky single attribute case (see Edwards, 1954, 1961; Edwards and Tversky, 1967; Lee, 1970). Only three studies analyzed the validity of models in the risky multiattribute case (VI)(Tverksy, 1967; v. Winterfeldt, 1971; Fischer, 1972). Few of these studies used the full potential of measurement theories, e.g., an analysis of their axioms, or application of formally justified assessment. Typically, some approximation method or approximate model was analyzed for its ability to predict global preference behavior or judgments. I do not know of

any descriptive applications of the bisymmetric or difference utility models.

Utility theory has also been applied as prescriptive theory of preferences. Any of the utility models described in this report could be used for modeling and measuring preferences to aid decision makers in evaluating or decision-making. Keeney and Raiffa (1975) list a number of cases in which utility theory--in particular expected utility measurement for single and multiattributed cases--has been applied for solving real world decision problems. There are extensive applications of simple expected utility models to business decision problems (see Schlaifer, 1969; Brown, 1970; Matheson, 1970), and a few applications of simple multiattribute models (see Keeney and Raiffa, 1975; Edwards et al., 1975, and Fischer and Edwards, 1973). In only a very few cases the full potential of utility theory (tests of independence assumptions and use of formally justified assessment) has been exploited in these applications. Applications of models for time variable cases and group preferences is still very limited. There has been no application of difference, bisymmetric, or conjoint measurement models for solving real decision problems.

Although some utility theories have been used in modeling preferences --descriptively in the laboratory or normatively in the real world--the general impression is that the application of utility theory is limited to a few classes of decision problems and utility models. If one would measure the usefulness of utility theory in terms of its present use in descriptive or normative modeling, the picture would indeed be rather gloomy. A better criterion--and the one that will be adopted in the following pages--is how utility theory could be used. Although the following discussion will concentrate on the potential use of utility theories as prescriptive models, many of the arguments apply to their descriptive use as well. Prescriptive use of utility theory means any application of axioms, measurement procedures or other formal ways of thinking in the framework of utility theory when dealing with a formal analysis of real world decision or evaluation problems. Ultimately, the usefulness of utility theory in this context is its ability to aid the decision maker (group, organization) in making decisions.

Utility theory as formulated in the preceding sections could have two related functions in such a formal analysis of decision problems:

1. To aid in eliciting a model that best fits the decision problems and that at the same time represents the decision maker's tastes, values, and preferences;
2. To elicit the appropriate numbers which are needed to permit the calculation of utilities.

The question, "how useful is utility theory?" then boils down to the question, "how well is utility theory equipped to solve these two tasks?"

Although the report will go into much detail discussing this question, the general conclusion may be stated here already: utility theory offers a fine conceptual framework for thinking about problem structure, model forms, and possible places where models can go wrong. With the help of utility theory, decision analysts may improve their model building process substantially. However, in the actual construction of scales, utility theory has little to offer, and analysts usually will have to use their own intuition and expertise when it comes to assessment of utilities. The reasons are, among others, that the procedures to assess utilities within the formally justified framework are often clumsy, complicated, and difficult to understand. They do not allow for errors, are time consuming, and they most often involve imaginary questions that are hard to think about.

In the following, first the use of utility theory to elicit models for preferences will be discussed. Then the use of utility theory for eliciting actual scales will be criticized, and, finally, some critical remarks will be made regarding the scope of utility theory in general.

Use of utility theory for model elicitation -- How can utility theory be used to help building models of preferences that can improve the decision making or evaluation process? First of all, if an analyst wants to build a model of preferences, he has to ask himself: utility for whom, and for what entity? These two questions identify the class of models that are applicable to the special decision problem. Utility theory does little for the analyst here but stress the importance of this question. Different models apply to different decision making and choice entities. But through classification schemes like the one presented in this report, some structure is provided for the pro-

cess of identifying the decision maker and the decision problem. Unfortunately, the most realistic and complex problems have not been modeled yet. If an analyst finds, for example, a decision problem in which groups make decisions that affect many with uncertain and multiattributed outcomes, he will find no model applicable to this case. He is left to his own expertise in either putting various models together (e.g., expected utility models, multiattributed models, and so forth), or to make up models himself.

The next step an analyst has to go through if he wants to make use of any utility model, is to define the problem into the format in which the models are formulated. He has to define value dimensions (attributes) of outcomes, the individuals affected by the outcomes, the time periods of payoffs, and all the uncertain contingencies under which outcomes will be received. Although utility theory itself offers little help in this important task of structuring the problem, it obviously makes beneficial contributions by defining what elements are required in a formal treatment of decision problems. Decision trees (Raiffa, 1968) and goal hierarchies (Manheim and Hall, 1967) are among the tools that have grown from utility theory.

Utility theory may also help to structure a decision problem in a way that suggests simple models and simple assessment, thus making the judgmental tasks easier. For example, if one wants to use additive evaluation models in a multiattribute context, one would try to structure the problem in a way that suggests preferential independence of attributes. This is often not very difficult to do. For example, when evaluating apartments on the attributes "distance from the office" and "transportation facilities", preferential independence will certainly be violated. One could get around that problem by redefining a compound attribute "accessibility of the office". In other words, often the concepts of utility theory help to identify possible problem areas in modeling early in the structuring phase. (Interestingly enough, this function of utility theory may defeat the purpose of sophisticated modeling itself; if in 99% of all evaluation problems it is possible to find a structure that suggests additive or other simple models, why bother developing much more complicated alternatives?)

After the structuring process--which really is a process of making the decision problem accessible to modeling--the decision analyst can begin to ask

questions to select appropriate models for solving the decision problem. At this point, utility theory may have its most beneficial applications. Of course, the analyst has little hope of ever identifying the one and only optimal model, but utility theory offers technical advice on how to go about eliminating models that are obviously wrong. The essence of this technical advice is contained in the various model assumptions and axioms.

Consideration of these assumptions, in thought experiments and discussions between analyst and decision maker may identify those assumptions that are clearly unacceptable to the decision maker or that do not fit the decision problem. The goal of this search and check procedure is to find the set of strongest assumptions that can still be accepted and to identify a model that meets this set of assumptions. v. Winterfeldt and Fischer (1975) recently presented a hierarchical structure of model assumptions for the riskless and risky multiattributed cases that a decision analyst can run through in order to eliminate models, and select a model form that appears acceptable.

Naturally, nothing guarantees that the model thus identified is an optimal one. Exploration of alternative models is good practice in decision analysis. Often one may want to make intentional modeling errors by using a model whose assumptions are violated in order to buy simpler assessment methods. In this case, of course, extensive sensitivity analyses should assure that these model violations won't lead to preposterous results.

After thought experiments (which are nothing else but imaginary examples of the formal model assumptions), the analyst can go further and use explicit model tests, by checking practically if some simple model implications or numerical assumptions hold. Again, utility theory provides such procedures. For example, if one wants to check the validity of the multiplicative expected utility model in risky multiattribute evaluation problems, one could check if certainty equivalents in one attribute are independent of constant values in other attributes. Of course, one has to take care that the task for the assessor is not so complex that it leads to errors simply because of information overload problems. Similar checks in a conjoint measurement framework may take the form of assessing single attribute utility functions conditional on various constant values in other attributes to see if the form of the single attribute function depends on these values.

Although these types of axiomatic check procedures are built into a utility theory, they are often formulated in such an invidious mathematical language that it is hard to recover their meaning from the model. It is crucial for such use of utility theory that the meaning of these assumptions is clear, otherwise there is little an analyst can communicate to a decision maker to check if the model assumptions are acceptable. For example, a decision maker may be baffled when asked "do your evaluations of risky multiattributed alternatives depend solely on their marginal probability distributions?" Unfortunately, utility theorists have often given in to mathematical sophistication in their model formulations, which usually hides the meaning of model assumptions. Besides formulating their assumptions in a rather mathematical fashion, utility theorists also try to make them appear very weak. A large amount of measurement theoretic effort is devoted to weakening assumptions while still retaining strong results. Fishburn and Keeney (1974), for example, show that with the help of some riskless assumptions (preferential independence) it is possible to derive powerful risky independence assumptions (utility independence). The problem with this kind of reasoning is that the "stronger" assumptions must still be valid if the "weaker" assumptions hold. One may be tempted to accept the validity of weak assumptions and forget about their strong implications. Although it may be misleading to ask the decision maker whether or not he would accept the full implication of all assumptions together (e.g., an additive expected utility model), it is equally erroneous to ask for acceptance of the weakest assumptions. To state it in mathematical language: testing the lemmas may often be more insightful than testing the axioms or theorems.

After structuring and after selection of a candidate model, the analyst can make use of utility theory in a final way when actually assessing the utility functions within the framework of the model he selected. In an extension of the numerical tests to select models, he can now build into the model construction various <u>consistency checks</u> that are suggested by the model form. In the multiattribute framework, for example, he may assess single attribute utility functions and predict some simple implications of the additive model with these functions. In the risky multiattribute case, he could assess more than one single attribute utility function to see if the shape of these functions

varies systematically with the conditioning values in other attributes. There are many ways of performing such consistency checks, and although utility theory provides guidance through the model forms, it remains an art to design such checks intelligently.

A final point that should be mentioned is that utility theory is silent about the errors that occur when one applies an inappropriate model to a decision problem. The model may be inappropriate for two reasons: first, because it makes assumptions that are not justified, usually because they are too strong. The reverse kind of mistake is also possible, namely applying a model--although its assumptions are uncontroversial--which is not strong enough. Generally, the weaker the model, the more complicated the procedures to assess utilities; and the more complicated the assessment procedures, the more likely assessment errors will occur. Deterministic utility theory--the only class of utility models that has any potential in real world applications--cannot say anything about this very real trade-off between model weakness and assessment complexity, since it does not acknowledge error. It is uncontroversial that a weak order model is always an appropriate normative model for decision making, since it only says that choice entities are to be ordered transitively. However, that statement helps little in model construction. Assessment errors that can occur, for example, in complex indifference curve assessment and trade-offs may be much more substantial than the errors that occur from assuming a model form that violates some of the decision maker's preferences. Probabilistic models appear to be no way out of this dilemma; although they acknowledge the possibility of error, they do not provide any mechanisms to realistically model and assess preferences in real world decision problems.

The message of this section addressed to the practitioner of decision analysis is this: yes, utility theory can help you structuring decision problems, and eliminating inappropriate models. What you have to do yourself is to translate often rather mathematical axioms into behaviorally meaningful and testable assumptions. And even after you go through the process of model elimination and selection, you will still have to make up your mind about the possible trade-offs between assessment error and modeling error. The message to utility theorists is this: models are useful, the more meaningful and real their assump-

sumptions and implication are. So it may be useful to give up some mathematically sophisticated formulations in favor of more clumsy but realistic ones, and it may be more useful to formulate all model assumptions and implications in their full strength rather than hide the strength of the model behind seemingly weak axioms.

Use of utility theory for scaling -- As mentioned before, assessment is really the weak point of utility theories. If the models are useful for defining appropriate model forms, they seldom have produced elicitation methods that are very attractive in applied settings. An attractive assessment procedure would be one that is, among other things,

1. simple,
2. error free,
3. time and cost efficient,
4. realistic,

while still being in agreement with the modeled preferences. In general, the theoretically feasible assessment procedures do not rate very well on these criteria, when compared with the direct estimation methods that were discussed earlier as approximation procedures. An exception may be difference and bisection assessment, because of their close resemblance to direct estimation procedures.

Let us try to substantiate the claim that the formally justified methods for eliciting utilities do not fare very well in application. Really, we are talking about only five different methods:

1. indifference curves (weak orders)
2. difference standard sequences (difference measurement)
3. bisection standard sequences (bisymmetric measurement)
4. dual standard sequences (conjoint measurement)
5. indifference lotteries (expected utility measurement).

All these formally justified assessment procedures require, in one form or another, indifference judgments. Usually, the assessor manipulates one entity, or one variable aspect of a choice entity (a probability, a single attribute value, etc.) to match that choice entity against a standard. This method

is well known in psychophysics as the method of adjustment (Torgerson, 1958). Depending on the model form, these matches or indifference judgments involve more or less complex choice entities. Indifference curve procedures can involve choice entities that vary simultaneously on many attributes, while standard sequence procedures in conjoint measurement involve choice entities that vary only on two attributes at a time. Indifference lottery procedures typically involve only one variable (a probability, or a single attribute value), but in multiattribute models rather complex matches have to be established for rescaling, or assessing risk transformations (see p. 43).

If the choice entities that are to be matched vary on many value relevant aspects (as in indifference-curve procedures), the matching task can be very <u>complicated</u> for the assessor. But even if choice entities are simple, such as gambles for money, indifference procedures may still be quite complex. Systematic reliability studies are missing, but Davidson et al., (1957) give some indication of inconsistencies in very simple indifference judgments. In addition, psychophysical studies on indifference judgments in a variety of tasks show that there are systematic biases in indifference judgments such as constant error, etc. (Torgerson, 1958).

Clearly, any assessment procedure will produce some amount of <u>error</u>. The point is that indifference procedures may increase the amount of error by asking unusual questions about complex choice alternatives. The more complex the choice alternatives that are to be matched, the larger the error will be. Besides this type of random error due to the procedure, another measurement error is introduced in these assessment techniques. This error results from approximation of utilities through interpolation or curve fitting. Interpolation of utilities is necessary, if the measurement procedure did not provide the utilities for all the choice entities under consideration, but rather of a few points in that area. This will almost always be the case in the three standard sequence procedures, since, by their logic, they identify utilities generically without concern for the decision problem. Standard sequences determine the utilities of a well defined subset of choice entities, but this subset may be very different from the choice entities that are the subject of the decision problem. In this case, the utilities of the points of interest will have to

be approximated. Also, if one constructs indifference curves, the only way to avoid interpolation is to begin with the choice entities of interest and trade them off into one attribute. Otherwise, indifference curves will have to be approximated and utilities for points not located on the indifference curves will have to be estimated. The only procedure which provides for a straightforward utility assessment of <u>all</u> choice entities under study is the indifference lottery procedure in its variable probability version, where the choice entity of interest is fixed and probabilities are varied to generate a match between this choice entity and a gamble for two reference choice entities (see p. 40).

One could, of course, construct utilities with any of the five procedures that are finely graded to reduce approximation errors. But such a process may turn out to be very <u>time consuming</u> and <u>inefficient</u>. For example, in a simple riskless multiattribute evaluation problem involving 10 attributes and only 10 steps in each attribute more than 100 such indifference judgments would have to be made (including the judgments necessary for rescaling), if one accepts the additive model, and a much larger number would be required to achieve an equally fine grid if the additive model fails. In complex models like the multilinear model, just the indifference judgments required for rescaling can go into the hundreds. With 10 attributes, for example, the multilinear model 4b in the risky multiattributed case XI requires the assessment of 1022 scaling constants (see Keeney and Raiffa, 1975). This may be too much effort when the task is, for example, to compare three or four choice alternatives.

All five assessment procedures involve indifference judgments about <u>imaginary</u> choice entities that are not attainable in the decision problem. This fact is obvious for the indifference curve procedure and the standard sequence procedure. Since the assessor manipulates one variable aspect of the choice entities, he has to think about choices that, in reality, do not exist for him. The same is true for indifference lotteries, since the assessor has to think about "reference" outcomes that are usually not attainable, and about probabilities of events that have nothing to do with his decision problem. This lack of realism in assessment may produce quite serious judgmental problems. Often, the analyst can formulate the assessment task in a more realistic way, but he

will never get around the problem that indifference methods are by definition imaginary. There is another aspect to the problem of realism of assessment that may prove even more difficult to overcome. Some of the choice entities in utility assessment may not only be imaginary, but they may not even be conceivable. For example, when evaluating apartments with a conjoint measurement model, one may have to construct utility functions over the attributes "rent" and "size". The assessor may have to make indifference judgments about apartments with a very large size and very small rent, a rather unlikely combination which he may not be able to think about.

So much for the arguments that the formally justified indifference procedures are too complicated, produce too much error, are time consuming and unrealistic; some more so than others. In general, it appears that the rarely used difference and bisection procedures score better on these applied criteria than dual standard sequences, indifference curves, and indifference lottery procedures. But if one wants to find simple, quick, realistic methods that produce little error, one will have to look outside of the realm of theoretically feasible methods. Probably the most reasonable methods of this sort are magnitude estimation methods such as direct rating, direct judgment of utility differences, direct ratio assessment of weights, direct assessment of probabilities and utilities, etc. Clearly, they are uncomplicated. All the assessor has to do is to quantify his judgment on a numerical scale. They also are quick and more realistic since they need only be applied to the choice entities that are under study in the decision problem. If the number of choice entities is large, the assessment task may still be considerable, but seldom as large as in assessing the full utility function. Of course, the question of error remains. To use the approximation methods, one has to make sure that they produce utilities that are interpretable within the theoretical model in which they are to be applied. Some such arguments can be found in the section on behavioral similarities between assessment procedures in this report. Still, these direct estimation methods are approximation methods, and errors will be made by not reproducing the ideal utility function that would be assessed if the assessor could overcome all the cognitive problems in theoretically feasible methods. On the other hand, possible sources of error are reduced by assessing only the utili-

ties for the choice entities that are to be evaluated. Therefore, no interpolation is necessary. Also, by reducing complexity, one may reduce error. Without any experimental evidence, conclusions are hard to draw. But it appears that there is at least a reasonable trade-off between quick, simple, and realistic direct estimation methods that are not formally justified in the model context and the somewhat clumsy feasible indifference methods. This trade-off calls for experimentaiton.

A concluding perspective -- Utility theory is a collection of models and assessment procedures to measure utilities of various types of choice entites, for many different kinds of decision makers, groups, or organizations, and for numerous decision problems. Hopefully, this report has shown that utility theory offers a large number of models and assessment procedures and that it has many possible areas of application. While the preceding sections emphasized the wealth of utility models and assessment procedures and their potential use in application, these last few paragraphs will point out some limits of utility theory.

If one inspects the progress of the mathematical treatments on utility measurement over the last few years, certain trends become obvious:

1. Within one and the same decision problem or evaluation paradigm more and more sophisticated models are developed that generalize previous model forms. A typical example of this trend are the last five years of modeling risky multiattributed preferences. Starting with the basic expected utility model, more and more general decomposition forms were added to the original additive and multiplicative forms.

2. Existing models are transferred relatively intact to different cases that have similar formal characteristics. Recent models for risky group and time preferences, for example, are simple reinterpretations of the expected utility decomposition models which were developed for risky multiattributed cases.

3. Some of utility theory becomes increasingly removed from its areas of application by weakening assumptions or formulating them in a mathematically elegant, but often unintelligible form.

On the other hand, some of the most interesting and vital practical modeling problems still lie at the periphery of utility theory. Some of these problems are:

1. Group decision making;
2. Errors in measurement (in particular possible trade-offs between error in assessment and error in modeling);
3. Basis for measurement with simple judgmental assessment methods.

Anybody who is interested in the application of utility theory (either as descriptive or normative theory) is concerned about the problem of _real_ preferences and looks for theories that are based on _real_ decision, _real_ judgments, and _real_ decision makers. Much of the recent research suggests that utility theory is more involved with its own formalities than with these real properties of preferences. Maybe utility theory could become more useful if theorists begin to take judgments and preferences with all these real properties more seriously than the mathematics of the models that intend to represent them.

REFERENCES

Arrow, K.J. Social Choice and Individual Values New York: Wiley & Sons, 1951.

Becker, G.M.. DeGroot, M.H., and Marschak, Y. Stochastic models of choice behavior. Behavioral Science, 1963, 8, 41-55.

Boyd, D.W. A methodology for analyzing decision problems involving complex preference assessment. Stanford Research Institute, Menlo Park, Ca., 1970.

Brown, R.V. Do managers find decision theory useful? Harvard Business Review, 1970, 48, 78-89.

Brown, R.V., Kahr, A.S., and Peterson, C. Decision Analysis for the Manager. New York: Holt, Rinehart, and Winston, 1974.

Coombs, C.H. A review of the mathematical psychology of risk and risk taking. Michigan Mathematical Psychology Program, MMPP No. 72-6, University of Michigan, Ann Arbor, Mi., 1972.

Coombs, C.H., Bezembinder, T.G., and Goode, F.M. Testing expectation theories of decision making without measuring utility or subjective probability. Journal of Mathematical Psychology, 1967, 4, 73-102.

Coombs, C.H. and Huang, L.C. Polynomial psychophysics of risk. Journal of Mathematical Psychology, 1970, 7, 317-338.

Davidson, D., Suppes, P., and Siegel, S. Decision Making: An Experimental Approach. Stanford: Stanford University Press, 1957.

Dawes, R.M. and Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Edwards, W. The theory of decision making. Psychological Bulletin, 1954, 51, 380-417.

Edwards, W. Behavioral decision theory. Annual Review of Psychology, 1961, 12, 473-498.

Edwards, W. Social utilities. The Engineering Economist, Summer Symposium Series, VI, 1971.

Edwards, W., Guttentag, M., and Snapper, K. Effective evaluation: A decision theoretic approach. In E.L. Streuning and M. Guttentag (eds.) Handbook

of Evaluation Research, Vol. I. Beverly Hills, Ca.: Sage Publications, 1975 (in press).

Edwards, W. and Tversky, A. (eds.) Behavioral Decision Theory. Middlesex, England: Penguin Books, 1967.

Einhorn, H.J. and Hogarth, R.M. Unit weighting schemes for decision making. Organizational Behavior and Human Performance, 1975, 13, 171-192.

Farquhar, P.H. Fractional hypercube decompositions of multiattribute utility functions. Technical Report, No. 222, Dept. of Operations Research, Cornell University, Ithaca, N.Y., 1974 (a).

Farquhar, P.H. Pyramid and semicube decompositions of multiattribute utility functions. Rand Research Memorandum, No. RM-5323, 1974 (b).

Fischer, G.W. Four methods for assessing multiattribute utilities: An Experimental evaluation. Technical Report No. 037230-6-T, Engineering Psychology Laboratory, University of Michigan, Ann Arbor, Mi., 1972.

Fischer, G.W. Experimental applications of multiattribute utility models. In D. Wendt and C.A.J. Vlek (eds.) Utility, Probability, and Human Decision Making, Dordrecht, Holland: Reidel, 1975.

Fischer, G.W. and Edwards, W. Technological aids for inference, evaluation, and decision making: A review of research and experience. Technical Report, Engineering Psychology Laboratory, University of Michigan, Ann Arbor, Mi., 1973.

Fishburn, P.C. Methods for estimating additive utilities. Management Science, 1967, 13, 435-453.

Fishburn, P.C. Utility Theory for Decision Making. New York: Wiley & Sons, 1970.

Fishburn, P.C. Interval representation for interval orders and semi-orders. Journal of Mathematical Psychology, 1973, 10, 91-105 (a).

Fishburn, P.C. The Theory of Social Choice. Princeton, New Jersey: Princeton University Press, 1973 (b).

Fishburn, P.C. Bernoullian utilities for multiple factor situations. In J.L. Cochrane and M. Zeleney (eds.) Multiple Criteria Decision Making. Columbia, South Carolina: University of South Carolina Press, 1973 (c).

Fishburn, P.C. von Neumann and Morgenstern utility functions on two attributes. Operations Research, 1974, 22, 35-45 (a).

Fishburn, P.C. Social choice functions. SIAM Review, 1974, 16, 63-8 (b).

Fishburn, P.C. Lexicographic orders, utilities, and decision rules: A survey. Management Science, 1974, 11, 1442-1471 (c).

Fishburn, P.C. Paradoxes of voting. The American Political Science Review, 1974, 537-546 (d).

Fishburn, P.C. Nondecomposable conjoint measurement for bisymmetric structures. Journal of Mathematical Psychology, 1975, 12, 75-89.

Fishburn, P.C. and Keeney, R.L. Seven independence concepts and continuous multiattribute utility functions. Journal of Mathematical Psychology, 1974, 11, 294-326.

Fuchs, L. Partially Ordered Algebraic Systems. Reading, Mass.: Addison-Wesley, 1963.

Galanter, E. The direct measurement of utility and subjective probability. American Journal of Psychology, 1962, 75, 208-220.

v. Helmholtz, H. Zählen und Messen erkenntsnistheoretisch betrachtet. In Philosophische Aufsätze. Eduard Zeller gewidmet. Leitzig: Fues's Verlag, 1887, 17-52.

Hölder, O. Die Axiome der Quantität und die Lehre vom Mass. Bericht der Verh. der Kgl. Saechsischen Gesellschaft Wiss., Mathematisch-Phys. Klasses, Leipzig, 1901, 53, 1-64.

Huang, L.C. The expected risk function. Michigan Mathematical Psychology Program, MMPP-71-6, University of Michigan, Ann Arbor, Mi., 1971.

Huber, G.P. Methods for quantifying subjective probabilities and multiattribute utilities. Decision Sciences, 1974, 430-458.

Huber, G.P., Daneshgar, R., and Ford, D.L. An empirical study of five utility models for predicting job preferences. Organizational Behavior and Human Performance, 1971, 6, 267-282.

Keeney, R.L. Quasi-separable utility functions. Naval Research Logistics Quarterly, 1968, 15, 551-565.

Keeney, R.L. Multidimensional utility assessment: Theory, assessment, and application. Technical Report No. 43, Operations Research Center, MIT, 1969.

Keeney, R.L. Multiplicative utility functions. Operations Research, 1974, 22, 22-34.

Keeney, R.L. A group preference axiomatization with cardinal utility. IIASA Research Memorandum, No. RM-75-47, Laxenburg, Austria, 1975.

Keeney, R.L. and Kirkwood, C.W. Group decision making using cardinal social welfare functions. Technical Report, No. 83, Operations Research Center, MIT, 1973.

Keeney, R.L. and Raiffa, H. Decision analysis with multiple conflicting objectives, preferences, and value trade-offs, 1975 (in press).

Koopman, B.O. The axioms and algebra of intuitive probability. Annals of Mathematics, 1940, 41, 269-292.

Koopmans, T.C. Stationary utility and impatience. Econometrica, 1960, 28, 287-309.

Koopmans, T.C., Diamond, P.A., and Williamson, R.E. Stationary utility and time perspective. Econometrica, 1964, 32, 82-100.

Krantz, D.H. Conjoint measurement: The Luce-Tukey axiomatization and some extensions. Journal of Mathematical Psychology, 1964, 1, 248-277.

Krantz, D.H., Luce, R.D., Suppes, P., and Tversky, A. Foundations of Measurement, Vol. 1, New York: Academic Press, 1971.

Krantz, D.H. and Tversky, A. Conjoint measurement analysis of composition rules in psychology. Psychological Review, 1971, 78, 151-169.

Lee, W. Decision Theory and Human Behavior. New York: Wiley & Sons, 1971.

Luce, R.D. Semiorders and a theory of utility discrimination. Econometrica, 1956, 24, 178-191.

Luce, R.D. Individual Choice Behavior. New York: Wiley & Sons, 1959.

Luce, R.D. and Krantz, D.H. Conditional expected utility. Econometrica, 1971, 39, 253-271.

Luce, R.D. and Raiffa, H. Games and Decisions. New York: Wiley & Sons, 1957.

Luce, R.D. and Suppes, P. Preference, utility, and subjective probability. In R.D. Luce, R.R. Bush, and E. Galanter (eds.) Handbook of Mathematical Psychology, Vol. III, New York: Wiley & Sons, 1965, 294-410.

Luce, R.D. and Tukey, J.W. Simultaneous conjoint measurement: A new type of fundamental measurement. Journal of Mathematical Psychology, 1964, 1, 1-27.

MacCrimmon, K.R. An overview of multiple criteria decision making. In J.L. Cochrane and M. Zeleney (eds.) Multiple Criteria Decision Making. Columbia, South Carolina: University of South Carolina Press, 1973.

MacCrimmon, K.R. and Siu, J.K. Making trade-offs. Working paper, No. 233, University of British Columbia, 1974.

MacCrimmon, K.R. and Toda, M. The experimental determination of indifference curves. The Review of Economic Studies, 1969, 36, 433-451.

Manheim, M.L. and Hall, F. Abstract representation of goals: A method for making decisions in complex problems. In Transportation: A Service. Proceedings of the Sesquientennial Forum, American Society of Mechanical Engineers, New York, 1967.

Matheson, J.E. Decision analysis practice: Examples and insights. Proceedings of the Fifth International Conference on Operations Research, Venice, 1969 (J. Lawrence, ed.). London: Tavistock Publications, 1970, 677-691.

Meyer, R.F. On the relationships between the utility of assets, utility of consumption, and investment strategy in an uncertain, but time invariant world. In OR 69: Proceedings of the Fifth International Conference on Operations Research (J. Lawrence, ed.). London: Tavistock, 1970.

von Heumann, J. and Morgenstern, O. Theory of Games and Economic Behavior. Princeton: Princeton University Press, 1947.

Pfanzagl, J. Theory of measurement. New York. Wiley & Sons, 1968.

Pollak, R.A. Additive von Neumann and Morgenstern utility functions. Econometrica, 1967, 35, 485-494.

Pollak, R.A. The risk independence axiom. Econometrica, 1973, 41, 35-40.

Pollard, A.B. A normative model for joint time/risk preference decision problems. Stanford Research Institute, Menlo Park, California, 1969.

Pollatsek, A. and Tversky, A. A theory of risk. Journal of Mathematical Psychology, 1970, 7, 540-553.

Rapoport, A. Dynamic programming models for multistage decision making. Journal of Mathematical Psychology, 1967, 4, 48-71.

Rapoport, A. Research paradigms for the study of dynamic decision behavior. In D. Wendt and C.A.J. Vlek (eds.) Utility, Probability, and Human Decision Making. Dordrecht, Holland: Reidel, 1975.

Raiffa, H. Decision Analysis. Reading, Mass.: Addison-Wesley, 1968.

Raiffa, H. Preferences for multiattribute alternatives. RM-5868, The Rand Corporation. Santa Monica, Ca., 1969.

Savage, L.J. The Foundations of Statistics. New York: Wiley & Sons, 1954.

Schlaifer, R. Analysis of Decisions Under Uncertainty. New York: McGraw-Hill, 1969.

Slovic, P. and Lichtenstein, S. Comparison of Bayesian and regression approaches to study of human information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Stevens, S.S. A scale for the measurement of a psychological magnitude: loudness. Psychological Review, 1936, 43, 403-416.

Stevens, S.S. Mathematics, measurement, and psychophysics. In S.S. Stevens (ed.) Handbook of Experimental Psychology. New York: Wiley & Sons, 1951.

Stevens, S.S. Psychophysics. New York: Wiley & Sons, 1975.

Torgerson, W.S. Theory and Methods of Scaling. New York: Wiley & Sons, 1958.

Tversky, A. Additivity, utility, and subjective probability. Journal of Mathematical Psychology, 1967, 4, 175-202.

Tversky, A. Intransitivities of preferences. Psychological Review, 1969, 76, 31-48.

Tversky, A. Choice by elimination. Journal of Mathematical Psychology, 1972, 1-27 (a).

Tversky, A. Elimination by aspects: A theory of choice. Psychological Review, 1972, 79, 281-299 (b).

Vinogradov, A.A. Ordered algebraic systems. In R.V. Gamkrelize (ed.) Progress in Mathematics, Vol. 5. New York: Plenum Press, 1969, 77-126.

v. Winterfeldt, D. Multiattribute utility theory: Theoretical background and an experimental validation. Unpubl. Diploma Thesis, University of Hamburg, 1971.

v. Winterfeldt, D. Comments on Jacquet-Lagreze's paper. In D. Wendt and C.A.J. Vlek (eds) Utility, Probability, and Human Decision Making. Dordrecht, Holland: Reidel, 1975.

v. Winterfeldt, D. and Edwards, W. Evaluation of complex stimuli using multi-attribute utility procedures. Technical Report, Engineering Psychology Laboratory, University of Michigan, Ann Arbor, Mi., 1973.

v. Winterfeldt, D. and Fischer, G.W. Multiattribute utility theory: Models and assessment procedures. In D. Wendt and C.A.J. Vlek (eds.) Utility, Probability, and Human Decision Making. Dordrecht, Holland: Reidel, 1975.

TECHNICAL REPORTS DISTRIBUTION LIST

Code 455

Director, Engineering Psychology (5 cys)
Programs, Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Defense Documentation Center (12 cys)
Cameron Station
Alexandria, VA 22314

Lt. Col. Henry L. Taylor, USAF
OAD(E&LS) ODDR&E
Pentagon, Rm. 3D129
Washington, D.C. 20301

CDR Richard Schlaff
Office of Assistant Secretary of
Defense (Intelligence)
Pentagon
Washington, D.C. 20301

Dr. Robert Young
Director, Human Resources Research
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Human Factors Plans, OP987P7
Office of the Chief of Naval
Operations
Department of the Navy
Washington, D.C. 20350

Personnel Logistics Plans, OP987P10
Office of the Chief of Naval
Operations
Department of the Navy
Washington, D.C. 20350

Assistant Chief for Technology
Code 200
Office of Navel Research
800 N. Quincy Street
Arlington, VA 22217

Dr. M. Bertin
Office of Naval Research
Scientific Liaison Group
American Embasy - Room A-407
APO San Francisco 96503

Office of Naval Research (6 cys)
International Programs
Code 102IP
800 North Quincy Street
Arlington, VA 22217

Fleet Analysis and Support Division,
Code 230
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Naval Analysis Programs, Code 431
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Operations Research Programs,
Code 434
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Statistics and Probability Program,
Code 436
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Information Systems Program,
Code 437
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Director, ONR Branch Office
Attn: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Director, ONR Branch Office
Attn: Psychologist
536 South Clark Street
Chicago, IL 60605

Director, ONR Branch Office
Attn: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Dir., Naval Research Laboratory (6 cys)
Technical Information Division
Code 2627
Washington, D.C. 20375

Mr. Arnold Rubinstein
Naval Material Command, NAVMAT 0344
Department of the Navy
Washington, D.C. 20360

Navy Personnel Research and
Development Center
Management Support Department,
Code 210
San Diego, CA 92152

Dr. Charles Gettys
Code 305
Navy Personnel Research and
Development Center
San Diego, CA 92152

Human Factors Dept., Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis & Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Mr. J. Barber
Headquarters DA, DAPE-PBR
Washington, D.C. 20546

Dr. Joseph Zeidner
Dir., Organization and Systems
Research Laboratory
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Edgar M. Johnson
Organizations & Systems Research Lab.
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
1400 Wilson Boulevard
Arlington, VA 22209

Chief, Human Engineering Division
Aerospace Medical Research Lab.
Wright-Patterson AFB, OH 45433

Dr. Donald A. Topmiller
Chief, Systems Effect. Branch
Human Engineering Division
Wright-Patterson AFB, OH 45433

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Journal Supplement Abstract (3 cys)
Service
American Psychological Association
1200 17th Street, N.W.
Washington, D.C. 20036

Dr. Eugene Galanter
Columbia University
Department of Psychology
New York, NY 10027

Dr. Bruce M. Ross
Catholic University
Department of Psychology
Washington, D.C. 20064

Dr. Robert R. Mackie
Human Factors Research, Inc.,
Santa Barbara Research Park
6780 Cortona Drive
Goleta, CA 93017

Mr. Alan J. Pesch
Eclectech Associates, Inc.
Post Office Box 179
North Stonington, CT 06359

Director, Human Factors Wing
Defense & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsville, Toronto, Ontario
Canada

Dr. A.D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
England

Dr. David Zaidel
Road Safety Centre Tecnion City
Haifa
Israel

Prof. Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
Federal Republic of Germany

Dr. Marshall A. Narva
Army Research Institute for
Behavioral & Social Sciences
1300 Wilson Boulevard, Room 236
Arlington, VA 22209

Major David Dianich
DSMS
Building 202
Fort Belvior, VA 22060

Dr. C. Kelly
Decisions & Designs, Inc.
Suite 600
7900 Westpark Drive
McLean, VA 22101

Mr. George Pugh
General Research Corp.
7655 Old Springhouse Road
McLean, VA 22101

Mr. Gary W. Irving
Integrated Sciences Corp.
1532 Third Street
Santa Monica, CA 90401

Dr. Paul Slovic
Oregon Research Institute
Post Office Box 3196
Eugene, Oregon 97403

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. A.C. Miller III
Stanford Research Institute
Decision Analysis Group
Menlo Park, CA 94025

Dr. R.A. Howard
Stanford University
Stanford, CA 94305

Dr. Lee Roy Beach
Department of Psychology
University of Washington
Seattle, Washington 98195

Mr. Frank Moses
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Major Robert G. Gough
Associate Professor
Department of Economics, Geography
and Management
USAF Academy, Colorado 80840

Dr. A.L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Social Science Research Institute
Research Reports

75-1 N. Miller and G. Maruyama. Ordinal Position and Peer Popularity. January, 1975.

75-2 G. Maruyama and N. Miller. Physical Attractiveness and Classroom Acceptance. January, 1975.

75-3 J.R. Newman, S. Kirby and A.W. McEachern. Drinking Drivers and Their Traffic Records. February, 1975.

75-4 Detlof von Winterfeldt and Ward Edwards. Error in Decision Analysis: How to Create the Possibility of Large Losses by Using Dominated Strategies. April, 1975.

75-5 Peter C. Gardiner and Ward Edwards. Public Values: Multi-Attribute Utility Measurement for Social Decision Making. (Forthcoming as a chapter in *Human Judgment and Decision Processes: Formal and Mathematical Approaches*, Steven Schwartz and Martin Kaplan (eds.), summer 1975.) May, 1975.

75-6 J. Buell, H. Kagiwada, and R. Kalaba. SID: A Fortran Program for System Identification. May, 1975.

75-7 J. Robert Newman. Assessing the Reliability and Validity of Multi-Attribute Utility Procedures: An Application of the Theory of Generalizability. July, 1975.

75-8 David A. Seaver, Detlof v. Winterfeldt, and Ward Edwards. Eliciting Subjective Probability Distributions on Continuous Variables. August, 1975.

75-9 Detlof von Winterfeldt. An Overview, Integration, and Evaluation of Utility Theory for Decision Analysis. August, 1975.